Euro. Jnl of Applied Mathematics (2012), *vol.* 23, *pp.* 593–610.
doi:10.1017/S0956792512000137

A statistical application of the quantile mechanics approach: MTM estimators for the parameters of t and gamma distributions

A. KLEEFELD[1] and V. BRAZAUSKAS[2]

[1] *Department I (Mathematics, Natural Sciences, Computer Science), Brandenburgische Technische Universität, P.O. Box 101344, 03013 Cottbus, Germany*
email: kleefeld@tu-cottbus.de
[2] *Department of Mathematical Sciences, University of Wisconsin – Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA*
email: vytaras@uwm.edu

(*Received 18 January 2011; revised 24 April 2012; accepted 25 April 2012; first published online 18 May 2012*)

In this paper, we revisit the quantile mechanics approach, which was introduced by Steinbrecher and Shaw (Steinbrecher, G. & Shaw, W. T. (2008) Quantile mechanics. *European. J. Appl. Math.* **19**, 87–112). Our objectives are (i) to derive the method of trimmed moments (MTM) estimators for the parameters of gamma and Student's t distributions, and (ii) to examine their large- and small-sample statistical properties. Since trimmed moments are defined through the quantile function of the distribution, quantile mechanics seems like a natural approach for achieving objective (i). To accomplish the second goal, we rely on the general large sample results for MTMs, which were established by Brazauskas *et al.* (Brazauskas, V., Jones, B. L. & Zitikis, R. (2009) Robust fitting of claim severity distributions and the method of trimmed moments. *J. Stat. Plan. Inference* **139**, 2028–2043), and then use Monte Carlo simulations to investigate small-sample behaviour of the newly derived estimators. We find that, unlike the maximum likelihood method, which usually yields fully efficient but non-robust estimators, the MTM estimators are robust and offer competitive trade-offs between robustness and efficiency. These properties are essential when one employs gamma or Student's t distributions in such outlier-prone areas as insurance and finance.

Key words: Point estimation (62F10); Asymptotic properties of estimators (62F12); Robustness and adaptive procedures (62F35); Method of trimmed moments (62F99); Computational problems in statistics (65C60)

1 Introduction

A plot based on *percentiles*, or what we now call a *quantile* function, was introduced by Galton [4]. Actually, according to Hald [6], many facts about quantiles were known before 1900. After the intial popularity, however, the use of quantiles for statistical modelling has been eclipsed by the likelihood-based techniques and partially by methods related to moments. Nonetheless, quantile statistical thinking has its own modern proponents who argued that a unification of the theory and practice of statistical methods of data modelling might be possible by a quantile perspective (see [5] and [15]). Other authors continue to make contributions to the field of quantile-based inference, one of which serves as motivation for this paper.

The quantile mechanics approach was recently introduced by Steinbrecher and Shaw [18]. The authors of that paper noticed that for standard probability distributions the quantile function satisfies a non-linear ordinary differential equation of the second order. Their proposal was to approximate its solution using a power series expansion. It was found that this kind of approximation is computationally tractable and more reliable than other well-established approaches, e.g. the Cornish–Fisher expansion (see [1]). For illustrative purposes, they considered the standard normal, Student's t, gamma and beta distributions. In addition, extensions of the quantile mechanics approach to more general situations (e.g. multivariate quantiles, quantile dynamics governed by stochastic differential equations) were also discussed. Among several venues of application the relevance of such results and tools to Monte Carlo simulations in financial risk management was emphasised.

Besides being used in financial applications, the aforementioned probability distributions and their variants (e.g. log-gamma, log-normal, log-t, log-folded-normal, log-folded-t) are commonly pursued for fitting insurance claims data. For various real-data examples in actuarial science, the reader can be referred to [2], [3] and [11]. The attractiveness of these parametric families for insurance modeling is two-fold: (i) parsimony, i.e. they have few parameters that have to be estimated from the data, and (ii) flexibility, i.e. they can capture the highly asymmetric and heavy-tailed nature of insurance data. It is not hard to imagine that similar qualities of a probability distribution may appeal to researchers working in quantitative risk management (see [13]), reliability engineering, econometrics, applied mathematics and statistics, among others. For a comprehensive survey of continuous univariate distributions, associated statistical inference, along with numerous areas of application, see [9].

In this paper, we revisit the quantile mechanics approach with the main objective to introduce robust and efficient estimators for the parameters of Student's t and gamma distributions. Robust and efficient estimation of parameters of a probability distribution is not a new area of mathematical statistics. Fundamental contributions to this field date back to the 1960s (see [7] and [8]), and the literature has been steadily growing ever since. For comprehensive treatment of robust statistics, one should consult a recent book by Marrona, Martin and Yohai [12]. Typically, robust and efficient estimators belong to one of the following three general classes of statistics: – L-, M- or R-statistics (see [16]). (Here: L stands for *linear* in the 'linear combinations of order statistics'; M stands for *maximum* in the 'maximum likelihood type statistics'; R stands for *ranks* in the 'statistics based on ranks'.) It is not uncommon, however, to have estimators that can be reformulated within more than one of these classes. On the other hand, despite the existing overlap, each type of these statistics has its own appeal. M-statistics, for example, are arguably the most amenable to generalisation and often lead to theoretically optimal procedures. R-statistics can be recast in the context of hypothesis testing and enjoy a close relationship with the broad field of non-parametric statistics. L-statistics are fairly simple computationally and have a straightforward interpretation in terms of quantiles. Moreover, the latter class is gaining popularity in the actuarial literature owing to its elegant treatment of various risk measures (see [10] and [14], and the references cited therein). Thus, in view of this discussion, we will pursue estimators that are based on L-statistics.

The *method of trimmed moments* (MTM), recently introduced by Brazauskas *et al.* [3], is a general parameter estimation method that falls within the class of L-statistics. It is computationally simple, it works like the classical method of moments (hence, it is easy to understand how it operates on the data) and is particularly effective for fitting location-scale families or their variants. The gamma and Student's t distributions are not proper location-scale families, since they both have a shape parameter that appears in the quantile function in a non-linear fashion (i.e. as an argument of gamma function). Therefore, for these distributions a reliable approximation of the quantile function is necessary. These considerations lead us to the quantile mechanics approach.

The rest of the paper is organised as follows. In Section 2, we derive a power series approximation of the quantile function for general versions of the Student's t and gamma distributions. In Section 3, we start with a brief review of the method of trimmed moments, then construct MTM estimators for the paramaters of gamma and Student's t distributions. Asymptotic properties of these estimators are also discussed in Section 3. Corresponding results for the maximum likelihood method are presented in Section 4. Further, in Section 5, we perform an extensive Monte Carlo simulations study and use it to investigate small-sample behaviour of the newly derived estimators. Concluding remarks are provided in Section 6.

2 Quantile mechanics

We start by first deriving an infinite power series for the quantile function of Student's t (in Section 2.1) and gamma (in Section 2.2) distributions, which is done by employing the quantile mechanics approach of Steinbrecher and Shaw [18]. The quantile functions will be used later to construct the method of trimmed moments estimators for the parameters of these distributions.

2.1 Student's t distribution

The probability density function (pdf) of a location-scale Student's t distribution is given by

$$f(x|\theta,\sigma,\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sigma\sqrt{\nu\pi}} \frac{1}{(1+\frac{1}{\nu}((x-\theta)/\sigma)^2)^{(\nu+1)/2}}, \qquad -\infty < x < \infty, \tag{2.1}$$

where $\theta \in \mathbb{R}$ is the location parameter, $\sigma > 0$ is the scale parameter and $\nu > 0$ is the shape parameter. In most statistical applications ν emerges as a positive integer and is called the degrees of freedom. In this paper, however, we will allow a more general definition of ν. Also, if $\nu = 1$, then (2.1) reduces to the pdf of Cauchy(θ,σ). If $\nu \to \infty$, then it converges to the pdf of normal(θ,σ). Note that Steinbrecher and Shaw [18] derived a power series expansion of the quantile function $F_{\mathrm{t},0}^{-1}$ for the standard case $f(x|0,1,\nu)$ and extensions to the general case is just given by $F_{\mathrm{t}}^{-1} = \theta + \sigma F_{\mathrm{t},0}^{-1}$. For the sake of completeness, we briefly review their derivation.

https://doi.org/10.1017/S0956792512000137 Published online by Cambridge University Press

To begin with, let $w(u)$ be the quantile function as a function of u, where $0 < u < 1$. With the relationship $\frac{dw}{du} = \frac{1}{f(w)}$ between the pdf and the quantile function, we have

$$\frac{dw}{du} = \sigma\sqrt{\nu\pi}\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\left(1+\frac{1}{\nu}((w-\theta)/\sigma)^2\right)^{(\nu+1)/2}. \tag{2.2}$$

Differentiating (2.2) with respect to u yields

$$\frac{d^2w}{du^2} = \frac{\nu+1}{\nu}\frac{1}{1+\frac{1}{\nu}((w-\theta)/\sigma)^2}\frac{w-\theta}{\sigma^2}\left(\frac{dw}{du}\right)^2,$$

which results in the ordinary differential equation (ODE) with centre condition

$$\begin{aligned}\sigma\left(1+\frac{1}{\nu}((w-\theta)/\sigma)^2\right)\frac{d^2w}{du^2} &= \left(1+\frac{1}{\nu}\right)\frac{w-\theta}{\sigma}\left(\frac{dw}{du}\right)^2,\\ w(1/2) &= \theta,\\ w'(1/2) &= \sigma\sqrt{\nu\pi}\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}.\end{aligned} \tag{2.3}$$

Applying the transformation

$$z = \sqrt{\nu\pi}\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}(u-1/2)$$

to the non-linear ODE (2.3) yields

$$\begin{aligned}\sigma\left(1+\frac{1}{\nu}((w-\theta)/\sigma)^2\right)\frac{d^2w}{dz^2} &= \left(1+\frac{1}{\nu}\right)\frac{w-\theta}{\sigma}\left(\frac{dw}{dz}\right)^2,\\ w(0) &= \theta,\\ w'(0) &= \sigma.\end{aligned} \tag{2.4}$$

Now, we assume that the solution of (2.4) is given by the infinite power series

$$w = \theta + \sigma\sum_{p=0}^{\infty} c_p z^{2p+1}. \tag{2.5}$$

Substituting this series into the ODE (2.4) and simplifying, yields

$$\begin{aligned}&\sigma^2\sum_{p=0}^{\infty}(2p+1)(2p)c_p z^{2p-1}\\ &= \sigma^2\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} c_k c_l c_m z^{2k+2l+2m+1}\left(\left(1+\frac{1}{\nu}\right)\frac{1}{\sigma}(2l+1)(2m+1)-\frac{1}{\nu\sigma^2}(2k+1)(2k)\right),\end{aligned}$$

https://doi.org/10.1017/S0956792512000137 Published online by Cambridge University Press

and we obtain the explicit cubic recurrence (see [18] for details)

$$(2p+1)(2p)c_p = \sum_{k=0}^{p-1}\sum_{l=0}^{p-k-1} c_k c_l c_{p-k-l-1}\left(\left(1+\frac{1}{\nu}\right)(2l+1)(2p-2k-2l-1)-\frac{1}{\nu}(2k+1)(2k)\right). \quad (2.6)$$

In the following we list some c_ps:

$$\begin{aligned}
c_0 &= 1,\\
c_1 &= \frac{1}{6}\frac{\nu+1}{\nu},\\
c_2 &= \frac{1}{120}\frac{(\nu+1)(7\nu+1)}{\nu^2},\\
c_3 &= \frac{1}{5040}\frac{(\nu+1)(127\nu^2+8\nu+1)}{\nu^3},\\
c_4 &= \frac{1}{362880}\frac{(\nu+1)(4369\nu^3-537\nu^2+135\nu+1)}{\nu^4},\\
c_5 &= \frac{1}{39916800}\frac{(\nu+1)(243649\nu^4-90488\nu^3+26238\nu^2-2504\nu+1)}{\nu^5}.
\end{aligned}$$

In summary, the quantile function for Student's t distribution is given by

$$F_{\mathrm{t}}^{-1}(u) = \theta + \sigma\sum_{p=0}^{\infty} c_p\left(\sqrt{\nu\pi}\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}(u-1/2)\right)^{2p+1}, \qquad 0<u<1, \quad (2.7)$$

where the coefficients c_p are given by (2.6).

2.2 Gamma distribution

The pdf of a two-parameter gamma distribution is given by

$$f(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}, \qquad x>0, \quad (2.8)$$

where $\beta > 0$ is the scale parameter and $\alpha > 0$ is the shape parameter. Note that a power series expansion of the quantile function $F_{\mathrm{g},0}^{-1}$ has been derived in [18] for the standard case $f(x|\alpha,1)$ and the extension to the general case is just $F_{\mathrm{g}}^{-1} = F_{\mathrm{g},0}^{-1}/\beta$ for $f(x|\alpha,\beta)$. For the sake of completeness, we briefly review their derivation.

Using the relation between pdf and quantile function, we obtain

$$\frac{dw}{du} = \frac{\Gamma(\alpha)}{\beta^\alpha}w^{1-\alpha}\mathrm{e}^{\beta w}. \quad (2.9)$$

Differentiating (2.9) with respect to u yields

$$\frac{d^2w}{du^2} = \left(\frac{dw}{du}\right)^2\left(\beta+\frac{1-\alpha}{w}\right),$$

which results in the ODE with left condition

$$w\frac{d^2w}{du^2} - (w\beta + 1 - \alpha)\left(\frac{dw}{du}\right)^2 = 0, \tag{2.10}$$
$$w(0) = 0,$$
$$w(u) \sim \frac{1}{\beta}\left[u\Gamma(\alpha+1)\right]^{1/\alpha} \quad \text{as} \quad u \to 0.$$

Applying the transformation

$$z = \left[u\Gamma(\alpha+1)\right]^{1/\alpha}$$

to the non-linear ODE (2.10) yields

$$w\left(\frac{d^2w}{dz^2} + \frac{1-\alpha}{z}\frac{dw}{dz}\right) - (w\beta + 1 - \alpha)\left(\frac{dw}{dz}\right)^2 = 0, \tag{2.11}$$
$$w(0) = 0,$$
$$w'(0) = \frac{1}{\beta}.$$

Assume that the solution of (2.11) is given by the infinite power series

$$w(z) = \frac{1}{\beta}\sum_{p=1}^{\infty} d_p z^p \qquad \text{with} \quad d_1 = 1. \tag{2.12}$$

Substituting the series into the ODE (2.11) yields

$$\begin{aligned} p(p+\alpha)d_{p+1} = &\sum_{k=1}^{p}\sum_{l=1}^{p-k+1} d_k d_l d_{p-k-l+2}\, l(p-k-l+2) \\ &- \Delta(p)\sum_{k=2}^{p} d_k d_{p-k+2}\, k\left[k - \alpha - (1-\alpha)(p+2-k)\right], \end{aligned} \tag{2.13}$$

where $\Delta(p) = 0$ if $p < 2$ and $\Delta(p) = 1$ if $p \geqslant 2$. In the following we list some d_ps:

$$d_1 = 1,$$
$$d_2 = \frac{1}{1+\alpha},$$
$$d_3 = \frac{1}{2}\frac{5+3\alpha}{(1+\alpha)^2(2+\alpha)},$$
$$d_4 = \frac{1}{3}\frac{31+33\alpha+8\alpha^2}{(1+\alpha)^3(2+\alpha)(3+\alpha)},$$
$$d_5 = \frac{1}{4}\frac{28888+1179\alpha^3+125\alpha^4+5661\alpha+3971\alpha^2}{(1+\alpha)^4(2+\alpha)^2(3+\alpha)(4+\alpha)}.$$

In summary, the quantile function for a gamma distribution is given by

$$F_{\mathrm{g}}^{-1}(u) = \frac{1}{\beta}\sum_{p=1}^{\infty} d_p\left(\left[u\Gamma(\alpha+1)\right]^{1/\alpha}\right)^p, \qquad 0 < u < 1, \tag{2.14}$$

where the coefficients d_p are given by (2.13).

3 MTM estimation

Throughout this paper we will consider a sample of n independent and identically distributed random variables, $X_1, \ldots, X_n$, from a cumulative distribution function (cdf) denoted by F. We assume that it is given in a parametric form and the k unknown parameters are denoted by $\theta_1, \ldots, \theta_k$. Also, the order statistics of $X_1, \ldots, X_n$ will be denoted by $X_{1:n} \leqslant \cdots \leqslant X_{n:n}$. Further, as presented by Brazauskas *et al.* [3], the MTM procedure is similar to the method of moments approach. The inherent difference is that we match population and sample-*trimmed* moments instead of matching ordinary moments. That is,

- first, we compute k sample-trimmed moments,

$$\hat{\mu}_j = \frac{1}{n - m_n(j) - m_n^*(j)} \sum_{i=m_n(j)+1}^{n-m_n^*(j)} h_j(X_{i:n}), \qquad j = 1, \ldots, k,$$

  where $m_n(j)$ and $m_n^*(j)$ are integers satisfying $0 \leqslant m_n(j) < n - m_n^*(j) \leqslant n$, and $m_n(j)/n \to a_j$ and $m_n^*(j)/n \to b_j$ as $n \to \infty$. The proportions a_j and b_j are chosen by the researcher as well as the functions $h_j : \mathbb{R} \to \mathbb{R}$.

- Second, the corresponding population-trimmed moments

$$\mu_j := \mu_j(\theta_1, \ldots, \theta_k) = \frac{1}{1 - a_j - b_j} \int_{a_j}^{1-b_j} h_j(F^{-1}(u))\, \mathrm{d}u, \qquad j = 1, \ldots, k,$$

  are derived with F^{-1} denoting the quantile function of F.

- Finally, the population and sample-trimmed moments are equated, and the resulting system of equations

$$\mu_j(\theta_1, \ldots, \theta_k) = \hat{\mu}_j, \qquad j = 1, \ldots, k,$$

  is solved with respect to $\theta_1, \ldots, \theta_k$. The solution of the system of equations is denoted by

$$\hat{\theta}_j = g_j(\hat{\mu}_1, \ldots, \hat{\mu}_k), \qquad j = 1, \ldots, k,$$

  and is, by definition, the MTM estimator of parameters $\theta_1, \ldots, \theta_k$.

The MTM estimator $(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ is *consistent* and *asymptotically normal* ($\mathscr{AN}$) with mean $(\theta_1, \ldots, \theta_k)$ and the covariance matrix $n^{-1}\,\mathbf{D\Sigma D}'$,

$$\left(\widehat{\theta}_1, \ldots, \widehat{\theta}_k\right) \sim \mathscr{AN}\left((\theta_1, \ldots, \theta_k), n^{-1}\,\mathbf{D\Sigma D}'\right),$$

https://doi.org/10.1017/S0956792512000137 Published online by Cambridge University Press

where $\mathbf{D} = [d_{ij}]_{i,j=1}^{k}$ is the Jacobian of transformations $g_1, \ldots, g_k$ evaluated at $(\mu_1, \ldots, \mu_k)$, that is $d_{ij} = \partial g_i / \partial \widehat{\mu}_j |_{(\mu_1, \ldots, \mu_k)}$, and $\mathbf{\Sigma} := [\sigma_{ij}^2]_{i,j=1}^{k}$ is a covariance matrix with

$$\sigma_{ij}^2 = \frac{1}{(1-a_i-b_i)(1-a_j-b_j)} \int_{a_i}^{1-b_i} \int_{a_j}^{1-b_j} (\min\{u,v\} - uv)\, \mathrm{d}h_j(F^{-1}(v))\, \mathrm{d}h_i(F^{-1}(u)).$$

For further technical details and examples, see [3].

3.1 Student's t distribution

To obtain the MTM estimator of θ, σ and v, we first compute the following three sample-trimmed moments:

$$\hat{\mu}_1 = \frac{1}{n - m_n(1) - m_n^*(1)} \sum_{i=m_n(1)+1}^{n-m_n^*(1)} Y_{i:n},$$

$$\hat{\mu}_2 = \frac{1}{n - m_n(1) - m_n^*(1)} \sum_{i=m_n(1)+1}^{n-m_n^*(1)} Y_{i:n}^2,$$

$$\hat{\mu}_3 = \frac{1}{n - m_n(2) - m_n^*(2)} \sum_{i=m_n(2)+1}^{n-m_n^*(2)} Y_{i:n}^2.$$

Note that by choosing $h_2(x) = h_3(x) = x^2$ we can ensure that the estimates of σ and v will be positive, which is desirable since these two parameters are positive. Next, we derive the three corresponding population-trimmed moments with the quantile function given by (2.7). We have

$$\begin{aligned} \mu_1 &= \frac{1}{1-a_1-b_1} \int_{a_1}^{1-b_1} F_{\mathrm{t}}^{-1}(u)\, \mathrm{d}u \\ &= \theta + \sigma \frac{1}{1-a_1-b_1} \int_{a_1}^{1-b_1} \sum_{p=0}^{\infty} c_p \left(\sqrt{v\pi} \frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)} (u - 1/2) \right)^{2p+1} \mathrm{d}u. \end{aligned} \tag{3.1}$$

With the definition

$$\lambda_{i,q} := \frac{1}{1-a_i-b_i} \int_{a_i}^{1-b_i} \left[\sum_{p=0}^{\infty} c_p \left(\sqrt{v\pi} \frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)} (u - 1/2) \right)^{2p+1} \right]^q \mathrm{d}u$$

we can write (3.1) as

$$\mu_1 = \theta + \sigma \lambda_{1,1}(v),$$

and similarly

$$\begin{aligned} \mu_2 &= \theta^2 + 2\theta\sigma\lambda_{1,1}(v) + \sigma^2\lambda_{1,2}(v), \\ \mu_3 &= \theta^2 + 2\theta\sigma\lambda_{2,1}(v) + \sigma^2\lambda_{2,2}(v). \end{aligned}$$

Equating $\hat{\mu}_1$ to μ_1, $\hat{\mu}_2$ to μ_2 and $\hat{\mu}_3$ to μ_3, and solving the resulting system of equations with respect to θ, σ and v yields the MTM estimator:

$$\hat{\theta}_{\text{MTM}} = \hat{\mu}_1 - \lambda_{1,1}(\hat{v}_{\text{MTM}})\hat{\sigma}_{\text{MTM}}, \tag{3.2}$$

$$\hat{\sigma}_{\text{MTM}} = \sqrt{\frac{\hat{\mu}_2 - \hat{\mu}_1^2}{\lambda_{1,2}(\hat{v}_{\text{MTM}}) - \lambda_{1,1}^2(\hat{v}_{\text{MTM}})}}, \tag{3.3}$$

$$\begin{aligned} \hat{\mu}_3 = {} & \hat{\mu}_1^2 + 2\hat{\mu}_1(\lambda_{2,1}(\hat{v}_{\text{MTM}}) - \lambda_{1,1}(\hat{v}_{\text{MTM}}))\sqrt{\frac{\hat{\mu}_2 - \hat{\mu}_1^2}{\lambda_{1,2}(\hat{v}_{\text{MTM}}) - \lambda_{1,1}^2(\hat{v}_{\text{MTM}})}} \\ & + \frac{\hat{\mu}_2 - \hat{\mu}_1^2}{\lambda_{1,2}(\hat{v}_{\text{MTM}}) - \lambda_{1,1}^2(\hat{v}_{\text{MTM}})} \\ & \times (\lambda_{1,1}^2(\hat{v}_{\text{MTM}}) - 2\lambda_{1,1}(\hat{v}_{\text{MTM}})\lambda_{2,1}(\hat{v}_{\text{MTM}}) + \lambda_{2,2}(\hat{v}_{\text{MTM}})). \end{aligned} \tag{3.4}$$

In summary, we first have to solve the non-linear equation (3.4) for $\hat{v}_{\text{MTM}}$. Then we can obtain $\hat{\sigma}_{\text{MTM}}$, defined by (3.3), and after that $\hat{\theta}_{\text{MTM}}$, defined by (3.2).

Note that by interchanging summation and integration we can simplify $\lambda_{i,1}$ and $\lambda_{i,2}$ to

$$\lambda_{i,1} = \frac{1}{1 - a_i - b_i} \sum_{p=0}^{\infty} \left\{ \frac{\Lambda^{2p+1}}{2p+2} \left[\left(\frac{1}{2} - b_i\right)^{2p+2} - \left(a_i - \frac{1}{2}\right)^{2p+2} \right] c_p \right\}, \tag{3.5}$$

$$\lambda_{i,2} = \frac{1}{1 - a_i - b_i} \sum_{p=0}^{\infty} \left\{ \frac{\Lambda^{2p+2}}{2p+3} \left[\left(\frac{1}{2} - b_i\right)^{2p+3} - \left(a_i - \frac{1}{2}\right)^{2p+3} \right] \sum_{k=0}^{p} c_k c_{p-k} \right\}, \tag{3.6}$$

where

$$\Lambda = \sqrt{v\pi} \frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}.$$

Finally, note that we can and have to approximate $\lambda_{1,1}(v)$, $\lambda_{1,2}(v)$, $\lambda_{2,1}(v)$ and $\lambda_{2,2}(v)$ by calculating a finite number of terms in the infinite series. Our numerical studies presented in Section 5 are based on a 50-term approximation, which guarantees a relative error of 10^{-5} for our considered MTMs (see Appendix for details).

3.2 Gamma distribution

To obtain the MTM estimator of α and β, we first compute the two sample-trimmed moments:

$$\hat{\mu}_1 = \frac{1}{n - m_n(1) - m_n^*(1)} \sum_{i=m_n(1)+1}^{n-m_n^*(1)} Y_{i:n},$$

$$\hat{\mu}_2 = \frac{1}{n - m_n(2) - m_n^*(2)} \sum_{i=m_n(2)+1}^{n-m_n^*(2)} Y_{i:n},$$

where the choice $h_1(x) = h_2(x) = x$ ensures that the estimates of $\alpha > 0$ and $\beta > 0$ are positive. Then we derive the two corresponding population-trimmed moments with the help of the quantile function given by (2.14). We have

$$\begin{aligned}\mu_1 &= \frac{1}{1-a_1-b_1}\int_{a_1}^{1-b_1} F_g^{-1}(u)\,\mathrm{d}u\\ &= \frac{1}{1-a_1-b_1}\int_{a_1}^{1-b_1} \frac{1}{\beta}\sum_{p=1}^{\infty} d_p\,(u\Gamma(\alpha+1))^{p/\alpha}\,\mathrm{d}u\\ &= \frac{1}{\beta}\frac{1}{1-a_1-b_1}\int_{a_1}^{1-b_1} \sum_{p=1}^{\infty} d_p\,(u\Gamma(\alpha+1))^{p/\alpha}\,\mathrm{d}u\\ &:= \frac{1}{\beta}\delta_1(\alpha)\end{aligned}$$

and similarly

$$\mu_2 = \frac{1}{\beta}\delta_2(\alpha).$$

Equating $\hat{\mu}_1$ to μ_1 and $\hat{\mu}_2$ to μ_2, and solving the resulting system of equations with respect to α and β yields the MTM estimator

$$\hat{\beta}_{\mathrm{MTM}} = \frac{\delta_1(\hat{\alpha}_{\mathrm{MTM}})}{\hat{\mu}_1}, \tag{3.7}$$

$$\frac{\delta_1(\hat{\alpha}_{\mathrm{MTM}})}{\delta_2(\hat{\alpha}_{\mathrm{MTM}})} = \frac{\hat{\mu}_1}{\hat{\mu}_2}. \tag{3.8}$$

In summary, we first have to solve the non-linear equation (3.8) for $\hat{\alpha}_{\mathrm{MTM}}$. Then we can obtain $\hat{\beta}_{\mathrm{MTM}}$, defined by (3.7).

Note that by interchanging summation and integration we can simplify δ_i to

$$\delta_i = \frac{1}{1-a_i-b_i}\sum_{p=1}^{\infty}\left\{\frac{\alpha\Gamma^{p/\alpha}(\alpha+1)}{p+\alpha}\left[(1-b_i)^{p/\alpha+1} - a_i^{p/\alpha+1}\right]d_p\right\}. \tag{3.9}$$

Finally, our numerical studies presented in Section 5 are based on a 50-term approximation for $\delta_1(\alpha)$ and $\delta_2(\alpha)$, which guarantees a relative error of 10^{-5} for our considered MTMs with the exception of $\delta_1(2.5)$ and $\delta_2(2.5)$, where we have to use a 90-term approximation (see Appendix for details).

4 Maximum likelihood estimation

Under some standard regularity conditions (see, e.g. [16, Section 4.2]), the maximum likelihood estimators (MLE) are consistent and asymptotically fully efficient. Since the gamma and Student's t distributions do satisfy those regularity conditions, in Section 5 we will use MLEs for the parameters of these two families as benchmark estimators. That is, we will monitor the performance of MTM estimators relative to that of corresponding MLEs. In the following two sections we briefly present the key facts of MLEs for Student's t and gamma distributions.

4.1 Student's t distribution

Given a sample $\vec{X} = (X_1, \ldots, X_n)$, the MLE estimator is found by (numerically) maximising the log-likelihood function

$$\log \mathscr{L}(\theta, \sigma, \nu \mid \vec{X}) = n \log \Gamma\left(\frac{\nu+1}{2}\right) - n \log \Gamma\left(\frac{\nu}{2}\right) - \frac{n}{2} \log \pi + \nu n \log \sigma - \frac{\nu+1}{2} \sum_{i=1}^{n} \log\left(\nu\sigma^2 + (X_i - \theta)^2\right) + \frac{\nu n}{2} \log \nu.$$

A straightforward calculation shows that

$$\left(\hat{\theta}_{\mathrm{MLE}}, \hat{\sigma}_{\mathrm{MLE}}, \hat{\nu}_{\mathrm{MLE}}\right) \sim \mathscr{AN}\left((\theta, \sigma, \nu), n^{-1}\Sigma_0\right),$$

where Σ_0 is the inverse of the Fisher information matrix, which is given by

$$\Sigma_0^{-1} = \begin{pmatrix} \frac{\nu+1}{\sigma^2(\nu+3)} & 0 & 0 \\ 0 & \frac{2\nu}{\sigma^2(\nu+3)} & -\frac{2}{\sigma(\nu+1)(\nu+3)} \\ 0 & -\frac{2}{\sigma(\nu+1)(\nu+3)} & \Sigma_{0,33}(\nu) \end{pmatrix}$$

with

$$\Sigma_{0,33}(\nu) = -\frac{1}{4}\left[\psi\left(1, \frac{\nu+1}{2}\right) - \psi\left(1, \frac{\nu}{2}\right) + \frac{2(\nu+5)}{\nu(\nu+1)(\nu+3)}\right].$$

Here $\psi(1, \cdot)$ denotes the first polygamma function, which is the the first derivative of the digamma function $\psi(\cdot) = \Gamma'(\cdot)/\Gamma(\cdot)$.

4.2 Gamma distribution

Given a sample $\vec{X} = (X_1, \ldots, X_n)$, the MLE estimator for α and β is found by (numerically) maximising the log-likelihood function

$$\log \mathscr{L}(\alpha, \beta \mid \vec{X}) = n\alpha \log \beta - n \log \Gamma(\alpha) + (\alpha - 1) \sum_{i=1}^{n} \log X_i - \beta \sum_{i=1}^{n} X_i.$$

Similar to the previous case, a straightforward calculation shows that

$$(\hat{\alpha}_{\mathrm{MLE}}, \hat{\beta}_{\mathrm{MLE}}) \sim \mathscr{AN}((\alpha, \beta), n^{-1}\Sigma_1)$$

with

$$\Sigma_1 = \frac{1}{\alpha\psi(1, \alpha) - 1} \begin{pmatrix} \alpha & \beta \\ \beta & \psi(1, \alpha)\beta^2 \end{pmatrix},$$

where $\psi(1, \cdot)$ denotes the first polygamma function.

https://doi.org/10.1017/S0956792512000137 Published online by Cambridge University Press

5 Simulations

In this section, we use Monte Carlo simulations to augment the theoretical (large-sample) results with small-sample investigations. In particular, we are interested in the following questions. First, for how large the sample size n the asymptotic unbiasedness of MLE and MTM becomes valid? Second, what impact does the robustness or non-robustness of an estimator have on its bias and relative efficiency when the underlying model is contaminated?

To answer the questions of interest, we will use the following contamination model

$$F_{\varepsilon} = (1-\varepsilon)F_0 + \varepsilon\, G, \tag{5.1}$$

where F_0 is the central model, which in this paper will be assumed to be either a Student's t distribution with (θ, σ, v) or a gamma distribution with (α, β). Also, G is a contaminating distribution that generates observations violating the distributional assumption, and the contamination level ε represents the probability that a sample observation comes from the distribution G rather than F_0. Note that by choosing $\varepsilon = 0$ we can simulate a contamination-free scenario, which will allow us to answer the first question raised at the beginning of this section.

The general design of our Monte Carlo study is as follows. We generate 10,000 samples of size n using the underlying model (5.1). For each sample, we estimate the unknown parameters via MLE and MTM approaches. That is, for the Student's t distribution we estimate the location parameter θ, the scale σ and the shape v, and for the gamma distribution the shape α and the scale β. In the next step we calculate the average mean and relative efficiency (RE) of those 10,000 estimates. We repeat this process 10 times and 10 average means and 10 REs are again averaged and their standard deviations are reported. (Such repetitions is a quick way to assess standard errors of the estimated means and REs.) Thus, the reported standardised means are the average of 100,000 estimates divided by the true value of the parameter that we are trying to estimate. For computation of REs, we modify the definition of asymptotic relative efficiency (see, e.g. [16, Section 4.1]) by replacing all entries in the relevant covariance matrices with the corresponding mean-squared errors. Our Monte Carlo study is based on the following choice of simulation parameters:

- Distribution F_0: t with $\theta = 1$, $\sigma = 1$ and $v = 1, 2, 5$; and Gamma with $\beta = 1$ and $\alpha = 0.5, 1.0, 2.5$.
- Distribution G: t with $\theta = 1$, $\sigma = 1$, $v = 1$ (when F_0 is t); and Gamma with $\beta = 1$ and $\alpha = 0.2$ (when F_0 is Gamma).
- Level of contamination: $\varepsilon = 0, 0.01, 0.05, 0.10$.
- Sample size: $n = 100, 200, 500, 10,000$ (when F_0 is t); and $n = 25, 50, 100, 500$ (when F_0 is Gamma).
- Estimators: MLE and MTM with
  - $(a_1, b_1, a_2, b_2) = (0.05, 0.05, 0.05, 0.7)$, denoted mtm1;
  - $(a_1, b_1, a_2, b_2) = (0.05, 0.1, 0.3, 0.1)$, denoted mtm2;
  - $(a_1, b_1, a_2, b_2) = (0.1, 0.05, 0.1, 0.55)$, denoted mtm3;

Table 1. *Standardised mean of MLE and MTM estimators for selected values of* α *of the* Gamma *distribution with* $\beta = 1$*. The entries are mean values based on* 100,000 *simulated samples of size n*

α	Estimator	$n = 25$		$n = 50$		$n = 100$		$n = 500$		$n \to \infty$	
		α	β	α	β	α	β	α	β	α	β
0.5	MLE	1.19	1.10	1.09	1.04	1.04	1.02	1.01	1.00	1	1
	MTM1	1.19	1.13	1.01	1.02	1.04	1.02	1.01	1.00	1	1
	MTM2	1.72	1.54	0.93	0.91	1.05	1.03	1.01	1.01	1	1
	MTM3	1.22	1.16	1.04	1.05	1.04	1.03	1.01	1.00	1	1
	MTM4	1.65	1.48	1.38	1.26	1.05	1.03	1.01	1.01	1	1
	MTM5	1.06	1.04	0.84	0.87	1.06	1.03	1.01	1.01	1	1
1.0	MLE	1.16	1.11	1.07	1.05	1.04	1.02	1.01	1.00	1	1
	MTM1	1.14	1.13	1.00	1.01	1.03	1.02	1.00	1.00	1	1
	MTM2	1.36	1.32	0.93	0.93	1.04	1.03	1.01	1.00	1	1
	MTM3	1.18	1.15	1.03	1.04	1.03	1.02	1.00	1.00	1	1
	MTM4	1.34	1.29	1.23	1.19	1.04	1.03	1.01	1.00	1	1
	MTM5	1.00	1.01	0.90	0.91	1.05	1.03	1.01	1.01	1	1
2.5	MLE	1.14	1.12	1.07	1.06	1.03	1.03	1.01	1.01	1	1
	MTM1	1.09	1.10	0.97	0.98	0.99	1.00	0.97	0.98	1	1
	MTM2	1.21	1.22	0.94	0.94	1.02	1.02	0.99	1.00	1	1
	MTM3	1.11	1.11	0.99	1.01	0.98	0.99	0.96	0.97	1	1
	MTM4	1.20	1.20	1.15	1.15	1.01	1.02	0.99	0.99	1	1
	MTM5	0.97	0.98	0.91	0.92	1.03	1.03	1.00	1.00	1	1

Note: The range of standard errors for the simulated entries of α and β, respectively, are 0.15–4.90, 0.22–7.60 ($\times 10^{-3}$, for $\alpha = 0.5$); 0.18–2.20, 0.20–2.40 ($\times 10^{-3}$, for $\alpha = 1.0$); 0.18–2.00, 0.18–2.20 ($\times 10^{-3}$, for $\alpha = 2.5$).

- $(a_1, b_1, a_2, b_2) = (0.1, 0.1, 0.35, 0.1)$, denoted MTM4;
- $(a_1, b_1, a_2, b_2) = (0.15, 0.15, 0.3, 0.15)$, denoted MTM5.

The findings of the simulation study are summarized in Tables 1–4. The entries of the column $n \to \infty$ in Tables 1 and 2 follow from the theoretical results and not from simulations. They are included as target quantities.

Tables 1 and 2 provide the summarised information that relates to the first question of interest. The following conclusions emerge. First of all, for the gamma distribution (see Table 1), the bias of α and β estimators gets within several percentage points of the target for samples of size 100 or larger. Except for MTM4 and MTM5 estimators, one can probably claim that the same conclusion is true even for n as small as 50, but not for $n = 25$. And clearly, for $n = 500$ the bias of all estimators is practically eliminated. Second, for the Student's t distribution (see Table 2), we notice that all methods do a good job when estimating θ and σ, but things are less impressive for the shape parameter v. When $v = 1$, the results become satisfactory for $n \geqslant 200$ and for the majority of estimators even for $n \geqslant 100$. As the t distribution gets less heavy-tailed (when v increases), the bias of all methods vanishes at a much slower rate. For example, for $v = 5$ and $n = 500$, only the MLE is performing according to expectations. We did additional simulations for the case

Table 2. *Standardised mean of MLE and MTM estimators for selected values of v of the t distribution with $\theta = 1$ and $\sigma = 1$. The entries are mean values based on* 100,000 *simulated samples of size n*

		$n = 100$			$n = 200$			$n = 500$			$n \to \infty$		
v	Estimator	θ	σ	v	θ	σ	v	θ	σ	v			
1	MLE	1.00	1.00	1.02	1.00	1.00	1.01	1.00	1.00	1.00	1	1	1
	MTM1	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1	1	1
	MTM2	1.00	1.05	1.36	1.00	1.02	1.05	1.00	1.01	1.02	1	1	1
	MTM3	1.00	1.02	1.06	1.00	1.01	1.02	1.00	1.00	1.01	1	1	1
	MTM4	1.00	1.01	1.05	1.00	1.00	1.01	1.00	1.00	1.01	1	1	1
	MTM5	1.00	1.02	1.35	1.00	1.01	1.05	1.00	1.00	1.02	1	1	1
2	MLE	1.00	1.00	1.05	1.00	1.00	1.02	1.00	1.00	1.01	1	1	1
	MTM1	1.00	1.01	1.18	1.00	1.00	1.04	1.00	1.00	1.02	1	1	1
	MTM2	1.00	1.02	2.19	1.00	1.01	1.18	1.00	1.01	1.06	1	1	1
	MTM3	1.00	1.01	1.39	1.00	1.01	1.07	1.00	1.00	1.03	1	1	1
	MTM4	1.00	1.01	2.02	1.00	1.00	1.18	1.00	1.00	1.05	1	1	1
	MTM5	1.00	1.01	4.14	1.00	1.01	2.11	1.00	1.00	1.30	1	1	1
5	MLE	1.00	1.01	1.49	1.00	1.00	1.09	1.00	1.00	1.04	1	1	1
	MTM1	1.00	1.00	2.52	1.00	1.00	1.69	1.00	1.00	1.28	1	1	1
	MTM2	1.00	1.00	3.23	1.00	1.00	2.27	1.00	1.00	1.62	1	1	1
	MTM3	1.00	1.00	2.80	1.00	1.00	1.91	1.00	1.00	1.38	1	1	1
	MTM4	1.00	0.99	3.35	1.00	1.00	2.58	1.00	1.00	1.94	1	1	1
	MTM5	1.00	0.97	3.95	1.00	0.99	3.43	1.00	1.00	2.87	1	1	1

NOTE: The range of standard errors for the simulated entries of θ, σ, v, respectively, are 0.04–0.64, 0.02–0.84, 0.03–9.0 ($\times 10^{-3}$, for $v = 1$); 0.04–0.32, 0.03–0.48, 0.08–25.0 ($\times 10^{-3}$, for $v = 2$); 0.02–0.31, 0.03–0.53, 0.15–16.4 ($\times 10^{-3}$, for $v = 5$).

$n = 10,000$ and found that a few more MTMs were within a close range of the target. However, to get the MTM5 estimator for v within a reasonable range of the target, we need a very large sample size. This conclusion is not surprising at all because as v gets larger, the Student's t distributions differ only in the extreme tails and are essentially indistinguishable in the middle section. The closeness of t distributions when v is large can also be inferred from one of the classic 'rules of thumb' in statistics, which says that in most practical situations the t-test can be approximated by the z-test for $n \geqslant 30$ (i.e. the 'infinity' is only 30 observations away).

Next, we illustrate the behaviour of estimators under several data contamination scenarios. Contamination studies for the gamma distribution are summarised in Table 3, and the corresponding investigations for the Student's t distribution are reported in Table 4. In both cases we choose the central model so that the MTM performances were worst under the 'clean' data scenario (see Tables 1 and 2). This choice, however, did not prevent MTM estimators to outperform MLE under the non-clean scenarios. Indeed, by comparing the case $\varepsilon = 0$ with $\varepsilon = 0.01$ in Tables 3 and 4 we see that just 1% of outliers can completely erase the great advantage of the non-robust MLE method, which is true in terms of both the bias and the efficiency. Further, for the gamma distribution, we observe a joint drift away from the target by all estimators as ε increases. While the

https://doi.org/10.1017/S0956792512000137 Published online by Cambridge University Press

Table 3. *Standardised means and relative efficiencies of MLE and MTM estimators under the model* $F_\varepsilon = (1-\varepsilon)\,\text{Gamma}(\alpha = 2, \beta = 1) + \varepsilon\,\text{Gamma}(\alpha = 0.2, \beta = 1)$. *The entries are mean values based on* 100,000 *simulated samples of size* $n = 500$

		$\varepsilon = 0$		$\varepsilon = 0.01$		$\varepsilon = 0.05$		$\varepsilon = 0.10$	
Statistic	Estimator	α	β	α	β	α	β	α	β
MEAN	MLE	1.01	1.01	0.87	0.88	0.57	0.60	0.40	0.44
	MTM1	0.99	0.98	0.96	0.96	0.83	0.85	0.66	0.71
	MTM2	1.00	1.00	0.97	0.97	0.83	0.85	0.66	0.69
	MTM3	0.99	0.98	0.96	0.96	0.87	0.89	0.75	0.80
	MTM4	1.00	1.00	0.97	0.97	0.86	0.88	0.72	0.76
	MTM5	1.00	1.00	0.98	0.98	0.86	0.89	0.72	0.76
RE	MLE	0.97		0.42		0.21		0.19	
	MTM1	0.82		0.74		0.33		0.20	
	MTM2	0.76		0.70		0.31		0.19	
	MTM3	0.78		0.74		0.38		0.22	
	MTM4	0.72		0.69		0.34		0.20	
	MTM5	0.62		0.60		0.31		0.19	

MTMs exhibit smaller bias than the MLE, their relative efficiencies eventually become the same as that of the MLE. This is not unexpected for, as the level of data contamination reaches or exceeds MTMs' breakdown points, they also become uninformative. Finally, similar behaviour is observed for the Student's t distribution as well. In this case, however, one additional point is noteworthy. It seems that all methods of estimation are much more successful in separating θ and σ from v, which is evident from examining the bias part in Table 4. Indeed, only the MLE of σ underestimates its target by 2%, 6% and 8% as ε increases, whereas the MTMs of both θ and σ are completely unaffected by the contamination of data.

6 Conclusions

In this paper, we have developed MTM estimators for the parameters of a two-parameter gamma distribution and a three-parameter Student's t distribution. Large- and small-sample statistical properties of new estimators have been explored and compared to those of the maximum likelihood procedure. We have seen that MTM estimators are theoretically sound, computationally attractive and provide sufficient protection against various data contamination scenarios.

The MTM approach is based on matching trimmed moments of a population with their sample counterparts. Since trimmed moments are defined through the quantile function of the distribution, which, in our case, required reliable approximations, the quantile mechanics approach was essential to accomplish our objectives. We are fully aware that there remain unanswered questions. For example, (a) how to construct similar estimators for multivariate distributions? or (b) how to extend these ideas to dynamical systems that are governed by stochastic differential equations? These problems and related issues we intend to tackle in future projects.

Table 4. *Standardised means and relative efficiencies of MLE and MTM estimators under the model* $F_\varepsilon = (1-\varepsilon)\,\text{Student}(\theta = 1, \sigma = 1, v = 5) + \varepsilon\,\text{Student}(\theta = 1, \sigma = 1, v = 1)$. *The entries are mean values based on* 100,000 *simulated samples of size* $n = 10{,}000$

		$\varepsilon = 0$			$\varepsilon = 0.01$			$\varepsilon = 0.05$			$\varepsilon = 0.10$		
Statistic	Estimator	θ	σ	v	θ	σ	v	θ	σ	v	θ	σ	v
MEAN	MLE	1.00	1.00	1.00	1.00	0.98	0.87	1.00	0.94	0.63	1.00	0.92	0.51
	MTM1	1.00	1.00	1.01	1.00	1.00	0.98	1.00	1.00	0.86	1.00	1.00	0.75
	MTM2	1.00	1.00	1.02	1.00	1.00	0.99	1.00	1.00	0.87	1.00	1.00	0.75
	MTM3	1.00	1.00	1.02	1.00	1.00	0.98	1.00	1.00	0.87	1.00	1.00	0.75
	MTM4	1.00	1.00	1.04	1.00	1.00	1.01	1.00	1.00	0.89	1.00	1.00	0.78
	MTM5	1.00	1.00	1.16	1.00	1.00	1.11	1.00	1.00	0.95	1.00	1.00	0.82
RE	MLE		0.98			0.34			0.12			0.09	
	MTM1		0.37			0.39			0.17			0.09	
	MTM2		0.26			0.29			0.16			0.09	
	MTM3		0.32			0.34			0.17			0.09	
	MTM4		0.16			0.18			0.14			0.08	
	MTM5		0.03			0.03			0.05			0.06	

Acknowledgement

The authors thank anonymous referee for many helpful comments and suggestions.

Appendix

In this appendix we provide a justification for the 50-term approximations of $\lambda_{i,1}(v)$ and $\lambda_{i,2}(v)$, which are given by (3.5) and (3.6), respectively. First of all, recall that $\lambda_{i,q}(v)$, $q = 1, 2$, is a short-cut notation for

$$\frac{1}{1 - a_i - b_i} \int_{a_i}^{1-b_i} F_{\mathrm{t},0}^{-1}(u)\, \mathrm{d}u,$$

where parameters a_i and b_i are used to ensure the robustness of MTMs. They can also be interpreted as truncation points of the *central* power series. In our simulation studies the smallest trimming proportions are $a_1 = 0.05$ and $b_1 = 0.1$ (for MTM2). If we choose $a_i = 0$ and $b_i = 0$ (MTM simply becomes the standard method of moments), our results remain valid but the MTM estimators loose robustness and for $v \leqslant 2$ do not exist. Moreover, under such a scenario one has also to be concerned with an appropriate approximation of the tail (see [18] for a discussion on this issue). Hence, by choosing $a_i > 0$ and $b_i > 0$, we can achieve not only the statistical robustness of MTMs against data contamination but also the computational robustness as approximation of the quantile function in extreme tails (i.e. below a_i and above $1 - b_i$) is not necessary.

To augment the above discussion with quantitative comparisons, we employ *Mathematica* (see [19]) for computation of a closed form expression of $F_{\mathrm{t},0}^{-1}$, which involves the inverse of the regularized incomplete beta function (see [17, p. 45, Equation 24]). This allows us to evaluate $\lambda_{i,q}$, $q = 1, 2$, for the parameters of MTM2 and for $v = 1, 2, 5$, and

Table 5. *Comparison of exact solution with a series approximations with* $m = 20, 30, 40, 50$ *for Student's t distributions*

Parameter	Exact value	$m = 20$ rel. error	$m = 30$ rel. error	$m = 40$ rel. error	$m = 50$ rel. error
$\lambda_{1,1}(1)$	-2.549321802048_{-1}	3.7514_{-3}	3.2380_{-4}	3.0509_{-5}	3.0293_{-6}
$\lambda_{1,2}(1)$	$+2.516925438185_{+0}$	1.0779_{-2}	1.3066_{-3}	1.5880_{-4}	1.9306_{-5}
$\lambda_{2,1}(1)$	$+5.105816042167_{+0}$	1.0868_{-5}	8.6718_{-8}	7.6478_{-10}	7.0980_{-12}
$\lambda_{2,2}(1)$	$+1.018204609708_{+0}$	1.2541_{-4}	1.4459_{-6}	1.6670_{-8}	1.9178_{-10}
$\lambda_{1,1}(2)$	-1.365216100747_{-1}	9.0746_{-4}	6.5314_{-5}	5.3937_{-6}	4.8274_{-7}
$\lambda_{1,2}(2)$	$+1.024507974413_{+0}$	1.3028_{-3}	1.1235_{-4}	1.0614_{-5}	1.0560_{-6}
$\lambda_{2,1}(2)$	$+3.730166685487_{-1}$	2.0003_{-6}	1.3192_{-8}	1.0303_{-10}	2.4702_{-12}
$\lambda_{2,2}(2)$	$+5.371020314362_{-1}$	1.2700_{-5}	1.0202_{-7}	9.0307_{-10}	8.4768_{-12}
$\lambda_{1,1}(5)$	-1.008429908492_{-1}	3.0003_{-4}	1.9096_{-5}	1.4439_{-6}	1.2065_{-7}
$\lambda_{1,2}(5)$	$+6.769099301596_{-1}$	2.8021_{-4}	1.9568_{-5}	1.5842_{-6}	1.3949_{-7}
$\lambda_{2,1}(5)$	$+3.167162238617_{-1}$	5.8876_{-7}	3.4153_{-9}	2.4534_{-11}	7.6103_{-13}
$\lambda_{2,2}(5)$	$+3.891460670661_{-1}$	2.5993_{-6}	1.6747_{-8}	1.2637_{-10}	1.1546_{-12}

Table 6. *Comparison of the exact solution with a series approximation with* $m = 20, 30, 40, 50$ *for gamma distributions*

Parameter	Exact value	$m = 20$ rel. error	$m = 30$ rel. error	$m = 40$ rel. error	$m = 50$ rel. error
$\delta_1(0.2)$	$+8.519788165182_{-2}$	1.5724_{-7}	3.9247_{-10}	3.0211_{-12}	1.8413_{-12}
$\delta_2(0.2)$	$+1.205647654983_{-1}$	1.5741_{-7}	3.9128_{-10}	1.4034_{-12}	2.2239_{-13}
$\delta_1(0.5)$	$+3.297931967105_{-1}$	1.5027_{-4}	9.1733_{-6}	6.7179_{-7}	5.4638_{-8}
$\delta_2(0.5)$	$+4.551321416543_{-1}$	1.5425_{-4}	9.4167_{-6}	6.8962_{-7}	5.6086_{-8}
$\delta_1(1.0)$	$+7.864354357282_{-1}$	1.9251_{-3}	3.5186_{-4}	7.5935_{-5}	1.8049_{-5}
$\delta_2(1.0)$	$+1.032356585763_{+0}$	2.0776_{-3}	3.7973_{-4}	8.1948_{-5}	1.9478_{-5}
$\delta_1(2.5)$	$+2.223527770333_{+0}$	1.2218_{-2}	4.4929_{-3}	1.9088_{-3}	8.8251_{-4}
$\delta_2(2.5)$	$+2.703673227068_{+0}$	1.4235_{-2}	5.2346_{-3}	2.2240_{-3}	1.0282_{-3}

to compare them with (3.5) and (3.6). A summary of these calculations is presented in Table 6. (In the sequel, we abbreviate $const \cdot 10^{\pm p}$ by $const_{\pm p}$ for compactness.) It is save to conclude that a 50-term approximation suffices to guarantee a relative error of 10^{-5} for the MTM2. As expected, even better results are obtained for other MTM estimator for which higher trimming proportions were used. Admittedly, we are conservative here; we could have used less than 50 terms for $v = 2$ and 5.

A similar analysis was performed for the MTM2 estimator in the gamma model, i.e. for $\delta_1(\alpha)$ and $\delta_2(\alpha)$ given by (3.9), with $\alpha = 0.2, 0.5, 1.0, 2.5$. The results are presented in Table 6. Here we again conclude that a 50-term approximation guarantees a relative error of 10^{-5} for $\alpha = 0.2, 0.5$ and1.0. However, for $\alpha = 2.5$ we need a 90-term approximation to obtain the relative errors of 6.1850_{-5} for $\delta_1(2.5)$ and of 7.2060_{-5} for $\delta_2(2.5)$.

Note that in general (i.e. for arbitrary choice of parameters) we still cannot answer the question: How many terms should be chosen depending on the trimming proportions and the parameters to be estimated? This remains a topic of further research that could, for example, build upon the results of [17]. In that paper, explicit formulas for Student's t quantile function were derived assuming $v = 1, 2, 4$ and iterative schemes were provided for arbitrary even v and low odd v; the case of non-integer v was also considered.

References

[1] Abramowitz, M. & Stegun, I. A. (1972) *Handbook of Mathematical Functions*, Applied Mathematics Series, No. 55, National Bureau of Standards, Gaithersburg, MD.

[2] Brazauskas, V. & Kleefeld, A. (2011) Folded and log-folded-t distributions as models for insurance loss data. *Scand. Actuar. J.* **1**, 59–74.

[3] Brazauskas, V., Jones, B. L. & Zitikis, R. (2009) Robust fitting of claim severity distributions and the method of trimmed moments. *J. Stat. Plan. Inference* **139**, 2028–2043.

[4] Galton, F. (1875) Statistics by intercomparison with remarks on the law of frequency of error. *Phil. Mag.* **49**, 33–46.

[5] Gilchrist, W. (2000) *Statistical Modelling with Quantile Functions*, Chapman and Hall, Boca Raton, FL.

[6] Hald, A. (1998) *A History of Mathematical Statistics from 1750 to 1930*, Wiley, New York.

[7] Hampel, F. R. (1968) *Contributions to the Theory of Robust Estimation.* PhD. Dissertation, University of California, Berkeley, CA.

[8] Huber, P. J. (1964) Robust estimation of a location parameter. *Ann. Math. Stat.* **35**, 73–101.

[9] Johnson, N. L., Kotz, S. & Balakrishnan, N. (1994) *Continuous Univariate Distributions*, 2nd ed., Wiley, New York.

[10] Jones, B. L. & Zitikis, R. (2003), (2004) Empirical estimation of risk measures and related quantities (with discussion). *N. Am. Actuar. J.* **7**, **8**, 44–54, 114–117, 117–118.

[11] Klugman, S. A., Panjer, H. H. & Willmot, G. E. (2008) *Loss Models: From Data to Decisions*, 3rd ed., Wiley, New York.

[12] Maronna, R. A., Martin, D. R. & Yohai, V. J. (2006) *Robust Statistics: Theory and Methods*, Wiley, New York.

[13] McNeil, A. J. & Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton University Press, Princeton, NJ.

[14] Necir, A. & Meraghni, D. (2010) Estimating L-functionals for heavy-tailed distributions and application. *J. Probab. Stat.* Article ID 707146, 34 pp.

[15] Parzen, E. (2004) Quantile probability and statistical data modeling. *Stat. Sci.* **19**, 652–662.

[16] Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

[17] Shaw, W. T. (2006) Sampling Student's T distribution – use of the inverse cumulative distribution function. *J. Comput. Finance* **9**, 37–73.

[18] Steinbrecher, G. & Shaw, W. T. (2008) Quantile mechanics. *European J. Appl. Math.* **19**, 87–112.

[19] Wolfram, S. (2004) *The Mathematica Book*, 5th ed. Wolfram Media, Champaign, IL.