Comput Stat (2013) 28:809–829
DOI 10.1007/s00180-012-0330-4

ORIGINAL PAPER

On estimation of measurement error models with replication under heavy-tailed distributions

Jin-Guan Lin · Chun-Zheng Cao

Received: 3 January 2011 / Accepted: 16 April 2012 / Published online: 9 May 2012

Abstract Measurement error (errors-in-variables) models are frequently used in various scientific fields, such as engineering, medicine, chemistry, etc. In this work, we consider a new replicated structural measurement error model in which the replicated observations jointly follow scale mixtures of normal (SMN) distributions. Maximum likelihood estimates are computed via an EM type algorithm method. A closed expression is presented for the asymptotic covariance matrix of those estimators. The SMN measurement error model provides an appealing robust alternative to the usual model based on normal distributions. The results of simulation studies and a real data set analysis confirm the robustness of SMN measurement error model.

Keywords EM algorithm · Measurement error · Replicated measurement · Scale mixtures of normal distribution

1 Introduction

It is often assumed in classical statistical regression models that the covariates or explanatory variables are observed exactly. However, this assumption can be challenged since observed values of variables can often be considered error-prone measurements of the true covariates. Dramatically, ignoring such errors in covariates usually results in biased estimates of the regression coefficients. As a more realistic

J.-G. Lin (✉) · C.-Z. Cao
Department of Mathematics, Southeast University, Nanjing 210096, China
e-mail: jglin@seu.edu.cn

C.-Z. Cao
School of Math and Statistics, Nanjing University of Information Science and Technology,
Nanjing 210044, China

representation of classical regression models, measurement error (errors-in-variables) models assume that the independent variables are subject to error. A comprehensive study in measurement error models (MEM) can be found in Fuller (1987), Cheng and Van Ness (1999) and Carroll et al. (2006).

The linear MEM can be described as follows,

$$x_t = \xi_t + \delta_t, y_t = \eta_t + \varepsilon_t, \eta_t = \alpha + \beta\xi_t, \quad t = 1, \ldots, n, \tag{1}$$

where (x_t, y_t) are observed values, which are equal to true (latent) unobserved variables (ξ_t, η_t) plus the additive measurement errors $(\delta_t, \varepsilon_t)$. The latent variables ξ_t can be regarded as fixed unknown parameters (a functional model), or as independent, identically distributed random variables (a structural model). In this work, we only focus on the structural type model. As Reiersol (1950) showed, when normality is assumed, this model is not identifiable unless further information about the parameters can be found. This occurs because one can not establish a single relationship between the parameters of the distribution for (x_t, y_t) and the parameters of the model. A common solution of this problem is to assume some prior knowledge about the error variance (Cheng and Van Ness 1999). However, the non-identifiability issue will not appear in the replicated measurement error model (RMEM), in which the error variances can be estimated through the replicated data. Maximum likelihood (ML) estimation of the replicated structural model under normal distributions was solved by Chan and Mak (1979) and Isogawa (1985). Recently, Lin et al. (2004) derived an iterative EM algorithm to compute the ML estimators of the replicated model also under the normal distribution. However, the normality assumption is doubtful and suffers from a lack of robustness against outlying observations on the parameter estimates. Hence, it is very important to develop more robust model to fit the replicated measurement error data.

In this paper, we will assume scale mixtures of normal (SMN) distributions (Andrews and Mallows 1974) for the accommodation of extreme and outlying observations in RMEM. As the most important subclass of the elliptical symmetric distributions (Fang et al. 1990), the class of SMN distributions is a very flexible extension of normal distribution. Models based on SMN distributions will present more appealing robustness compared to the normal ones. Details about SMN distributions can be found in Andrews and Mallows (1974), Fang et al. (1990) and Lange and Sinsheimer (1993). Recently, SMN distributions have been applied to some special MEM. For example, Osorio et al. (2009) studied estimation and influence diagnostics for the Grubbs' model under SMN distributions; same issues on non-replicated MEM model based on SMN distributions are investigated by Lachos et al. (2011). Furthermore, scale mixtures of the skew-normal distributions have also been applied to MEM by Lachos et al. (2010). In this work, we discuss the ML estimation of RMEM, where the replicated observed values jointly follow SMN distributions. The hierarchical representation proposed by Pinheiro et al. (2001) is considered, which make it convenient to apply the EM algorithm for parameter estimation.

The rest of this paper is organized as follows. A brief sketch of SMN distributions is presented in Sect. 2. In Sect. 3, the replicated structural measurement error model with scale mixtures of normal distributions (SMN-RMEM) is defined, and an EM-type algorithm is applied to obtain the maximum likelihood estimates. A closed form expression is also obtained for the asymptotic covariance matrix of the ML estimators. Results of simulation studies are reported in Sect. 4. In Sect. 5, the CSFII (Continuing Survey of Food Intakes by Individuals) data (Thompson et al. 1992) is analyzed under the proposed SMN-RMEM. Some concluding remarks are given in the last section.

2 SMN distributions

SMN distributions, which play very important roles in statistical modeling, can be defined as the following m-dimensional random vector

$$Y = \boldsymbol{\mu} + \kappa^{1/2}(V)W, \tag{2}$$

where $\boldsymbol{\mu}$ is an m-dimensional location vector, W is an m-dimensional normal random vector with mean vector $\mathbf{0}$, and covariance matrix Σ, and $\kappa(\cdot)$ is a strictly positive weight function which is associated to the independent mixture variable V, a positive random variable with cumulative distribution function $H(v; \nu)$. It is easy to see from (2) that, the conditional distribution of Y given V is a multi-normal distribution, i.e., $Y|V = v \sim \mathrm{N}_m(\boldsymbol{\mu}, \kappa(v)\Sigma)$. Therefore, the marginal density function of Y takes the form as

$$f(\boldsymbol{y}) = |2\pi\Sigma|^{-1/2} \int_0^\infty \{\kappa(v)\}^{-m/2} \exp\{-\kappa^{-1}(v)u/2\} dH(v), \tag{3}$$

where $u = (\boldsymbol{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{y} - \boldsymbol{\mu})$ is the Mahalanobis distance. If Y has the form as (2) or has the density as (3), we will denote $Y \sim \mathrm{SMN}_m(\boldsymbol{\mu}, \Sigma; H)$. Owing to its special structure, the class of SMN distributions has similar properties to the normal distribution. With a suitable choice of $\kappa(\cdot)$ and the distribution function $H(\cdot; \nu)$, many heavy-tailed distributions can be generated, which are very useful for robust inference. Note that when $\kappa(\cdot) = 1$, the distribution of Y is just the normal one (N-SMN). Members of SMN distributions may be found, for instance, in Andrews and Mallows (1974). Here, three important examples are listed and will be applied in our study. We also compute the conditional expectation $\mathrm{E}[\kappa^{-1}(V)|Y]$ under the following distributions, which is helpful to carry out the EM algorithm.

(i) Multivariate Student-t distribution (T-SMN)

The multivariate Student-t distribution (Cornish 1954, Dunnett and Sobel 1954) with ν degrees of freedom, $t_m(\boldsymbol{\mu}, \Sigma; \nu)$, can be derived from the mixture structure (2), by taking $\kappa(v) = 1/v$ and $V \sim Gamma(\nu/2, \nu/2)$. The Cauchy distribution is obtained when $\nu = 1$, and one also gets the normal distribution when $\nu \to \infty$. In this case, the conditional expectation is

$$\mathrm{E}[\kappa^{-1}(V)|Y] = \frac{\nu + m}{\nu + u},$$

where u is the Mahalanobis distance we mentioned above.

(ii) Slash distribution (S-SMN)

To get the multivariate slash distribution (Rogers and Tukey 1972), denoted by $\mathrm{SL}_m(\boldsymbol{\mu}, \Sigma; \nu)$, one needs to take $\kappa(v) = 1/v$ and $V \sim Beta(\nu, 1)$, which has the density function as

$$h(v; \nu) = \nu v^{\nu-1}, \quad 0 < v \leqslant 1, \nu > 0.$$

It follows that

$$\mathrm{E}[\kappa^{-1}(V)|Y] = \left(\frac{2\nu + m}{u}\right) \frac{P_1(m/2 + \nu + 1, u/2)}{P_1(m/2 + \nu, u/2)},$$

where $P_x(a, b)$ denotes the cumulative distribution function of the $Gamma(a, b)$ distribution, i.e.,

$$P_x(a, b) = \frac{b^a}{\Gamma(a)} \int_0^x t^{a-1} e^{-bt} dt.$$

(iii) Contaminated normal distribution (C-SMN)

The multivariate contaminated normal distribution (Tukey 1960), $\mathrm{CN}_m(\boldsymbol{\mu}, \Sigma; \nu, \gamma)$, can be obtained from (2) by taking $\kappa(v) = 1/v$, and supposing V follows a discrete random probability function

$$h(v; \nu, \gamma) = \nu \mathrm{I}(v = \gamma) + (1 - \nu)\mathrm{I}(v = 1), \quad 0 \leqslant \nu \leqslant 1, 0 < \gamma \leqslant 1.$$

The density of Y has a mixture form as

$$f(\boldsymbol{y}) = |2\pi\Sigma|^{-1/2}[\nu\gamma^{m/2} e^{-\gamma u/2} + (1 - \nu)e^{-u/2}],$$

and the conditional expectation is

$$\mathrm{E}[\kappa^{-1}(V)|Y] = \frac{1 - \nu + \nu\gamma^{m/2+1} e^{(1-\gamma)u/2}}{1 - \nu + \nu\gamma^{m/2} e^{(1-\gamma)u/2}}.$$

3 Estimation of SMN-RMEM model

In this section, we will describe the SMN-RMEM model and investigate the EM algorithm and asymptotic covariance for the parameter estimators.

3.1 The SMN-RMEM model

Consider a bivariate random variable (ξ, η) satisfying a linear relationship $\eta = \alpha + \beta\xi$, in which ξ and η cannot be observed directly and we observe the values of $x = \xi + \delta$ and $y = \eta + \varepsilon$ with measurement errors δ and ε. For each ξ and η, p and q repeated observations $x_t^{(i)}, i = 1, \ldots, p$, $y_t^{(j)}, j = 1, \ldots, q$ are obtained, respectively. Then the SMN-RMEM model is given by

$$\begin{aligned} x_t^{(i)} &= \xi_t + \delta_t^{(i)}, \quad & i = 1, \ldots, p, \\ y_t^{(j)} &= \eta_t + \varepsilon_t^{(j)}, \quad & j = 1, \ldots, q, \\ \eta_t &= \alpha + \beta\xi_t, \quad & t = 1, \ldots, n, \end{aligned} \tag{4}$$

where $Z_t = (x_t^{(1)}, \ldots, x_t^{(p)}, y_t^{(1)}, \ldots, y_t^{(q)})^\top$ follows a SMN distribution, with the hierarchical structure form, which based on the suggestion of Pinheiro et al. (2001), as

$$\begin{aligned} Z_t|\xi_t, v_t &\overset{ind}{\sim} \mathrm{N}_m(\boldsymbol{a} + \boldsymbol{b}\xi_t, \kappa(v_t)\boldsymbol{D}(\boldsymbol{\phi})), \\ \xi_t|v_t &\overset{ind}{\sim} \mathrm{N}(\lambda, \kappa(v_t)\phi_\xi),\ v_t \overset{iid}{\sim} H(v; \nu),\ t = 1, \ldots, n, \end{aligned} \tag{5}$$

where $m = p + q$, $\boldsymbol{a} = (\overbrace{0, \ldots, 0}^{p}, \alpha\mathbf{1}_q^\top)^\top$, $\boldsymbol{b} = (\mathbf{1}_p^\top, \beta\mathbf{1}_q^\top)^\top$, in which $\mathbf{1}_p$ and $\mathbf{1}_q$ are, respectively, p- and q-dimensional vector with all elements are equal to 1, $\boldsymbol{\phi} = (\phi_\delta\mathbf{1}_p^\top, \phi_\varepsilon\mathbf{1}_q^\top)^\top$, and $\boldsymbol{D}(\cdot)$ denotes the diagonal transformation which transforms a vector to a diagonal matrix. In fact, it can be inferred from above that $Z_t \overset{iid}{\sim} \mathrm{SMN}_m(\boldsymbol{\mu}, \Sigma; H)$, where the location and scale parameter can be expressed as $\boldsymbol{\mu} = (\lambda\mathbf{1}_p^\top, (\alpha + \beta\lambda)\mathbf{1}_q^\top)^\top$, $\Sigma = \phi_\xi\boldsymbol{b}\boldsymbol{b}^\top + \boldsymbol{D}(\boldsymbol{\phi})$. Note that, model (4) is the same as model (1) when $p = q = 1$.

3.2 EM algorithm

As a special case, Lin et al. (2004) obtained ML estimators through an EM iteration for model (4) under normal distributions. Due to the hierarchical structure of (5), it is natural to also use the EM algorithm (Dempster et al. 1977, McLachlan and Krishnan 1997) to calculate the ML estimates of the parameters.

Let $\boldsymbol{\theta} = (\lambda, \alpha, \beta, \phi_\delta, \phi_\varepsilon, \phi_\xi)^\top$ be the parameter vector of model (4), and $\boldsymbol{\theta}^{(k)}$ denotes the estimates of $\boldsymbol{\theta}$ at the k-th iteration. Let $Z_c = (Z, \boldsymbol{\xi}, \boldsymbol{v})$ be the complete data set of model (4), where $Z = (Z_1^\top, \ldots, Z_n^\top)^\top$, $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)^\top$ and $\boldsymbol{v} = (v_1, \ldots, v_n)^\top$. It follows from (5) that the complete log-likelihood function associated with Z_c has the form as

$$\begin{aligned} l_c(\boldsymbol{\theta}|Z_c) \ &= -\frac{n}{2}\log\left(|\boldsymbol{D}(\boldsymbol{\phi})|\right) - \frac{1}{2}\sum_{t=1}^{n}\kappa^{-1}(v_t)(Z_t - \boldsymbol{a} - \boldsymbol{b}\xi_t)^\top\boldsymbol{D}^{-1}(\boldsymbol{\phi})(Z_t - \boldsymbol{a} - \boldsymbol{b}\xi_t) \\ &\quad -\frac{n}{2}\log(\phi_\xi) - \frac{1}{2\phi_\xi}\sum_{t=1}^{n}\kappa^{-1}(v_t)(\xi_t - \lambda)^2 + C, \end{aligned}$$

where $\log(|\boldsymbol{D}(\boldsymbol{\phi})|) = p\log(\phi_\delta) + q\log(\phi_\varepsilon)$, and C is a constant that is independent of $\boldsymbol{\theta}$. The EM algorithm is listed as follows.

E-step: Given the current estimate $\boldsymbol{\theta}^{(k)}$, the expected complete data log-likelihood function $\mathrm{E}[l_c(\boldsymbol{\theta}|Z_c)|\boldsymbol{\theta}^{(k)}, Z]$, also called the Q-function in Dempster et al. (1977), may be expressed as

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = -\frac{n}{2}\log\left(|\boldsymbol{D}(\boldsymbol{\phi})|\right) - \frac{1}{2}\sum_{t=1}^{n}\kappa_t^{(k)}(Z_t - \boldsymbol{a} - \boldsymbol{b}\xi_t^{(k)})^\top \boldsymbol{D}^{-1}(\boldsymbol{\phi})(Z_t - \boldsymbol{a} - \boldsymbol{b}\xi_t^{(k)})$$
$$- \frac{1}{2}\sum_{t=1}^{n}\tau^{(k)}\boldsymbol{b}^\top \boldsymbol{D}^{-1}(\boldsymbol{\phi})\boldsymbol{b} - \frac{n}{2}\log(\phi_\xi) - \frac{1}{2\phi_\xi}\sum_{t=1}^{n}\left[\kappa_t^{(k)}(\xi_t^{(k)} - \lambda)^2 + \tau^{(k)}\right].$$

where $\kappa_t^{(k)} = \mathrm{E}[\kappa^{-1}(v_t)|\boldsymbol{\theta}^{(k)}, Z_t]$, $\tau^{(k)} = \phi_\xi^{(k)}/S^{(k)}$ with $S^{(k)} = 1 + \phi_\xi^{(k)}\boldsymbol{b}^{(k)\top}\boldsymbol{D}^{-1}(\boldsymbol{\phi}^{(k)})\boldsymbol{b}^{(k)}$, and $\xi_t^{(k)} = \lambda^{(k)} + \tau^{(k)}\boldsymbol{b}^{(k)\top}\boldsymbol{D}^{-1}(\boldsymbol{\phi}^{(k)})(Z_t - \boldsymbol{a}^{(k)} - \boldsymbol{b}^{(k)}\lambda^{(k)})$.

M-step: Based on the Q-function, a new parameter estimate $\boldsymbol{\theta}^{(k+1)}$ can be obtained by maximize $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$ with respect to $\boldsymbol{\theta}$. The details of these steps are described as follows.

Firstly, the regression coefficients are updated by

$$\alpha^{(k+1)} = (b_1^{(k)}a_{22}^{(k)} - b_2^{(k)}a_{12}^{(k)})\big/(a_{11}^{(k)}a_{22}^{(k)} - a_{12}^{2(k)}),$$
$$\beta^{(k+1)} = (b_2^{(k)}a_{11}^{(k)} - b_1^{(k)}a_{12}^{(k)})\big/(a_{11}^{(k)}a_{22}^{(k)} - a_{12}^{2(k)}),$$

where $a_{11}^{(k)} = \sum_{t=1}^{n}\kappa_t^{(k)}$, $a_{22}^{(k)} = n\tau^{(k)} + \sum_{t=1}^{n}\kappa_t^{(k)}\xi_t^{2(k)}$, $a_{12}^{(k)} = \sum_{t=1}^{n}\kappa_t^{(k)}\xi_t^{(k)}$, $b_1^{(k)} = \sum_{t=1}^{n}\kappa_t^{(k)}\overline{y}_t$, $b_2^{(k)} = \sum_{t=1}^{n}\kappa_t^{(k)}\xi_t^{(k)}\overline{y}_t$, and $\overline{y}_t = \sum_{j=1}^{q}y_t^{(j)}/q$.

Then, other parameters can be updated by the following equations:

$$\lambda^{(k+1)} = a_{12}^{(k)}/a_{11}^{(k)}, \quad \phi_\xi^{(k+1)} = \tau^{(k)} + \frac{1}{n}\sum_{t=1}^{n}\kappa_t^{(k)}(\xi_t^{(k)} - \lambda^{(k+1)})^2,$$

$$\phi_\varepsilon^{(k+1)} = \tau^{(k)}(\beta^{(k+1)})^2 + \sum_{t=1}^{n}\kappa_t^{(k)}\left[\sum_{j=1}^{q}(y_t^{(j)} - \alpha^{(k+1)} - \beta^{(k+1)}\xi_t^{(k)})^2\right],$$

$$\phi_\delta^{(k+1)} = \tau^{(k)} + \frac{1}{np}\sum_{t=1}^{n}\kappa_t^{(k)}\left[\sum_{i=1}^{p}(x_t^{(i)} - \xi_t^{(k)})^2\right].$$

Starting with a suitable initial vector value $\boldsymbol{\theta}^{(0)}$, the algorithm iterates between the E- and M-steps until it reaches convergence. To ensure it is positive, the inverse of matrix Σ used in each E-step is computed by a closed form (Harville 1997) as

$$\Sigma^{-1} = \boldsymbol{D}^{-1}(\boldsymbol{\phi}) - \boldsymbol{D}^{-1}(\boldsymbol{\phi})\boldsymbol{b}\boldsymbol{b}^\top\boldsymbol{D}^{-1}(\boldsymbol{\phi})/(\phi_\xi^{-1} + \boldsymbol{b}^\top\boldsymbol{D}^{-1}(\boldsymbol{\phi})\boldsymbol{b}).$$

3.3 The expected information matrix

Since the SMN distributions belong to the elliptical distribution class (Fang et al. 1990), the replicated observations Z_t of model (4) can also be regarded as following an elliptical distribution $\mathrm{EL}_m(\boldsymbol{\mu}, \Sigma, g)$, where $\boldsymbol{\mu}$ and Σ are the same as those in Sect. 3.1, and $g(\cdot) : \mathbb{R} \to [0, \infty)$ is called the density generator such that $\int_0^\infty g(u)du < \infty$. Hence, the density function of Z_t takes the form

$$f(Z_t) = |\Sigma|^{-1/2} g(u_t),\ t = 1, \ldots, n,$$

where $u_t = (Z_t - \boldsymbol{\mu})^\top \Sigma^{-1}(Z_t - \boldsymbol{\mu})$, and $g(u_t)$ can be expressed as

$$g(u_t) = (2\pi)^{-m/2} \int_0^\infty \kappa(v)^{-m/2} \exp\{-\kappa^{-1}(v)u_t/2\} dH(v).$$

The log-likelihood function for model (4) is given by

$$l(\boldsymbol{\theta}) = -\frac{n}{2}\log(|\Sigma|) + \sum_{t=1}^{n} \log\{g(u_t)\}, \tag{6}$$

By calculating the expectations of the second-order derivatives of (6), we obtain the Fisher information matrix of $\boldsymbol{\theta}$ as $\mathbf{I}(\boldsymbol{\theta}) = (\mathrm{I}_{ij})_{6\times 6}$, with

$$\begin{aligned}
\mathrm{I}_{ij} &= \frac{4n}{m} d_g \dot{\boldsymbol{\mu}}_i^\top \Sigma^{-1} \dot{\boldsymbol{\mu}}_j,\ i \leqslant 3,\ j \leqslant 3, \quad \text{and} \quad (i, j) \neq (3, 3);\\
\mathrm{I}_{ij} &= na\mathrm{tr}(\Sigma^{-1}\dot{\Sigma}_i \Sigma^{-1}\dot{\Sigma}_j) + nb\mathrm{tr}(\Sigma^{-1}\dot{\Sigma}_i)\mathrm{tr}(\Sigma^{-1}\dot{\Sigma}_j),\\
&\quad\ i \geqslant 3,\ j \geqslant 3, \quad \text{and} \quad (i, j) \neq (3, 3);\\
\mathrm{I}_{33} &= \frac{4n}{m} d_g \dot{\boldsymbol{\mu}}_3^\top \Sigma^{-1} \dot{\boldsymbol{\mu}}_3 + na\mathrm{tr}(\Sigma^{-1}\dot{\Sigma}_3 \Sigma^{-1}\dot{\Sigma}_3) + nb\mathrm{tr}^2(\Sigma^{-1}\dot{\Sigma}_3);\\
\mathrm{I}_{ij} &= 0,\ \text{for } i \leqslant 2,\ j \geqslant 4 \text{ or } i \geqslant 4,\ j \leqslant 2,
\end{aligned}$$

where $a = \frac{2f_g}{m(m+2)}$, $b = \frac{f_g}{m(m+2)} - \frac{1}{4}$, $f_g = \mathrm{E}\{W_g^2(u)u^2\}$, $d_g = \mathrm{E}\{W_g^2(u)u\}$, in which $W_g(u) = \frac{\dot{g}(u)}{g(u)}$ with $u = \boldsymbol{e}^\top \boldsymbol{e}$ and $\boldsymbol{e} \sim \mathrm{EL}_m(\mathbf{0}, I_m)$, $\dot{\boldsymbol{\mu}}_i = \partial \boldsymbol{\mu}/\partial\theta_i$ and $\dot{\Sigma}_i = \partial\Sigma/\partial\theta_i$.

Due to the similarity between the inference for elliptical models and normal models, it is reasonable to expect that under suitable regularity conditions, the approximate distribution of the ML estimator $\widehat{\boldsymbol{\theta}}$ in large samples is $\mathrm{N}_6(\boldsymbol{\theta}, \mathbf{I}^{-1}(\boldsymbol{\theta}))$. Hence, the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ can be estimated by $\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}})$.

4 Simulation study

In this section, we perform Monte-Carlo simulations to compare the performance of the ML estimators of RMEM under four different type SMN distributions: N-SMN, T-SMN, S-SMN and C-SMN. Note that all of the other three distributions have heavier tails than the normal one, which indicates that statistical models based on those distributions are more robust to outliers. The degrees of freedom are set as follows: $\nu = 4$ (for T-SMN), $\nu = 3$ (for S-SMN) and $\nu = 0.1$, $\gamma = 0.2$ (for C-SMN). Other parameters in model (4) are set as: $\lambda = 3$, $\alpha = 2$, $\beta = 1$, $\phi_\delta = 1$ and $\phi_\xi = 1$.

The values of ϕ_ε has two values for comparison: 1 and 0.2, corresponding to the ratio of the error variances $\phi_\varepsilon/\phi_\delta$ of 1 and 0.2, respectively. The replicated numbers of the observations are chosen as $p = 3$ and $q = 2$.

In order to maintain the general and consistent form of the estimators among different SMN distributions, the degrees of freedom for the SMN model will not be estimated together with the interested parameters. In the simulation study, the main motivation is to confirm the heavy-tailed model's robustness and accuracy. Hence, we selected some heavy-tailed distributions with fixed degrees of freedom. While in the application, to find the best distribution for the data, we will choose the degrees of freedom by some usual criterions.

4.1 The first simulation

It is well known that misspecification of model's distribution will lead to the biases of parameter estimates. In the first simulation, we want to show that the ML estimates based on the heavy-tailed RMEMs will be more accurate than normal ones, when the true distribution of the data is heavy-tailed. In this simulation, we independently generate 2000 random samples with sample sizes $n = 20, 50$, and 100 from model (4) under one of the four SMN distributions. Then, we compute the ML estimators of $\boldsymbol{\theta}$ through the EM algorithm under all the four SMN distributions, respectively. Tables 1, 2, 3 and 4 display both the simulated sample means and standard deviations (SD) of interesting parameters λ, α and β, under simulated datasets generated by four different SMN distributions, respectively.

Some valuable conclusions can be drawn from the simulation study. For each case, the bias and the SD values are almost smallest when the true distribution is used. As expected, the SD of all estimates become smaller as n increases and as the variance radio $\phi_\varepsilon/\phi_\delta$ decreases. The most important information of Tables 2, 3 and 4 is that, when one of the three heavy-tailed distributions is assumed, the estimators under the normal distribution are worst at all times, since their SD values are largest among all estimators based on the four distributions. It is confirmed that RMEM under heavy-tailed distributions are more effective than the normal one, even if the distribution we used is not the true distribution.

4.2 The second simulation

In the second simulation, we will compare the performance of the estimators based on different methods in the presence of outliers. Regression calibration (RC) is also considered in the simulation. RC is an important estimation method for the MEM which can correct biases of the naive estimators. Details of this method may be found, for instance, in Carroll et al. (2006). Note that, we use analysis of variance formulae to get the consistent estimates of λ and variance components ϕ_δ and ϕ_ξ. Then, the RC estimator is obtained based on the adjustment of regression on averages of the observed variables, which is the same as the adjusting ordinary least squares mentioned in Lin et al. (2004).

Table 1 Performance of estimators under N-SMN-RMEM datasets

n	SMN type	$\phi_\varepsilon/\phi_\delta = 1$						$\phi_\varepsilon/\phi_\delta = 0.2$					
		λ		α		β		λ		α		β	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
20	N-SMN	3.0047	0.2645	1.9268	0.8793	1.0259	0.2820	3.0022	0.2573	1.9559	0.5527	1.0177	0.1775
	T-SMN	3.0052	0.2779	1.9315	0.9252	1.0241	0.2961	3.0020	0.2742	1.9539	0.5669	1.0173	0.1818
	S-SMN	3.0054	0.2662	1.9277	0.8847	1.0258	0.2837	3.0034	0.2594	1.9521	0.5545	1.0177	0.1781
	C-SMN	3.0056	0.2674	1.9262	0.8980	1.0260	0.2885	3.0048	0.2601	1.9514	0.5626	1.0179	0.1806
50	N-SMN	3.0011	0.1669	1.9604	0.4914	1.0140	0.1589	3.0012	0.1632	1.9906	0.3301	1.0037	0.1049
	T-SMN	3.0010	0.1764	1.9532	0.5278	1.0156	0.1702	3.0018	0.1736	1.9907	0.3440	1.0037	0.1092
	S-SMN	3.0015	0.1681	1.9600	0.4996	1.0139	0.1615	3.0014	0.1646	1.9889	0.3331	1.0045	0.1058
	C-SMN	3.0018	0.1685	1.9597	0.5070	1.0141	0.1639	3.0013	0.1651	1.9888	0.3375	1.0046	0.1071
100	N-SMN	2.9990	0.1163	1.9864	0.3375	1.0048	0.1077	2.9988	0.1124	1.9938	0.2193	1.0023	0.0706
	T-SMN	2.9987	0.1221	1.9830	0.3559	1.0060	0.1136	2.9967	0.1178	1.9927	0.2320	1.0025	0.0745
	S-SMN	2.9991	0.1167	1.9863	0.3417	1.0050	0.1091	2.9978	0.1128	1.9933	0.2219	1.0023	0.0714
	C-SMN	2.9994	0.1170	1.9860	0.3464	1.0050	0.1108	2.9982	0.1131	1.9930	0.2247	1.0024	0.0724

The underlined values in this table are obtained from the true model, i.e., the N-SMN-RMEM

Table 2 Performance of estimators under T-SMN-RMEM datasets

n	SMN type	$\phi_\varepsilon/\phi_\delta = 1$						$\phi_\varepsilon/\phi_\delta = 0.2$					
		λ		α		β		λ		α		β	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
20	N-SMN	3.0045	0.3593	1.7977	1.4753	1.0629	0.4675	2.9916	0.3534	1.8840	0.7887	1.0387	0.2541
	T-SMN	3.0007	0.2863	1.8650	1.0012	1.0425	0.3214	2.9949	0.2816	1.9529	0.5845	1.0170	0.1878
	S-SMN	3.0019	0.3028	1.8618	1.0599	1.0428	0.3373	2.9937	0.2948	1.9399	0.6157	1.0210	0.1979
	C-SMN	3.0017	0.3060	1.8556	1.0924	1.0446	0.3484	2.9948	0.2948	1.9411	0.6291	1.0213	0.2025
50	N-SMN	2.9961	0.2306	1.9018	0.8555	1.0337	0.2800	2.9919	0.2327	1.9367	0.5408	1.0215	0.1767
	T-SMN	3.0006	0.1874	1.9617	0.5602	1.0128	0.1791	2.9968	0.1878	1.9805	0.3520	1.0065	0.1110
	S-SMN	3.0007	0.1951	1.9555	0.5823	1.0151	0.1861	2.9955	0.1950	1.9772	0.3663	1.0076	0.1157
	C-SMN	3.0012	0.1969	1.9453	0.6021	1.0183	0.1933	2.9958	0.1977	1.9700	0.3859	1.0100	0.1225
100	N-SMN	2.9965	0.1639	1.9565	0.6366	1.0140	0.2052	2.9938	0.1598	1.9807	0.3856	1.0075	0.1245
	T-SMN	2.9991	0.1248	1.9890	0.3735	1.0039	0.1205	2.9944	0.1244	1.9986	0.2464	1.0012	0.0786
	S-SMN	2.9965	0.1316	1.9870	0.3861	1.0046	0.1239	2.9936	0.1309	1.9976	0.2591	1.0016	0.0820
	C-SMN	2.9970	0.1325	1.9839	0.4157	1.0055	0.1333	2.9936	0.1315	1.9956	0.2708	1.0023	0.0861

The underlined values in this table are obtained from the true model, i.e., the T-SMN-RMEM

Table 3 Performance of estimators under S-SMN-RMEM datasets

n	SMN type	$\phi_\varepsilon/\phi_\delta = 1$						$\phi_\varepsilon/\phi_\delta = 0.2$					
		λ		α		β		λ		α		β	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
20	N-SMN	3.0054	0.3224	1.8732	0.9623	1.0431	0.3053	3.0078	0.3106	1.9513	0.5982	1.0153	0.1903
	T-SMN	3.0047	0.3232	1.8802	0.9505	1.0426	0.3006	3.0068	0.3098	1.9578	0.5785	1.0143	0.1840
	S-SMN	3.0041	0.3149	1.8866	0.9198	1.0399	0.2905	3.0061	0.3015	1.9581	0.5715	1.0138	0.1820
	C-SMN	3.0045	0.3177	1.8790	0.9394	1.0424	0.2968	3.0072	0.3030	1.9517	0.5845	1.0150	0.1865
50	N-SMN	2.9963	0.2020	1.9611	0.5383	1.0127	0.1714	2.9969	0.1982	1.9645	0.3694	1.0117	0.1168
	T-SMN	2.9972	0.2064	1.9666	0.5257	1.0115	0.1660	2.9976	0.1993	1.9747	0.3503	1.0087	0.1106
	S-SMN	2.9975	0.1973	1.9712	0.5094	1.0095	0.1605	2.9985	0.1933	1.9776	0.3450	1.0084	0.1088
	C-SMN	2.9973	0.1975	1.9689	0.5248	1.0103	0.1657	2.9974	0.1938	1.9733	0.3497	1.0088	0.1104
100	N-SMN	3.0044	0.1403	1.9709	0.3962	1.0109	0.1258	2.9976	0.1419	1.9875	0.2526	1.0043	0.0813
	T-SMN	3.0032	0.1389	1.9718	0.3723	1.0104	0.1167	2.9984	0.1445	1.9916	0.2384	1.0029	0.0764
	S-SMN	3.0030	0.1357	1.9749	0.3631	1.0087	0.1141	2.9986	0.1394	1.9933	0.2336	1.0024	0.0747
	C-SMN	3.0041	0.1364	1.9740	0.3669	1.0097	0.1154	2.9978	0.1398	1.9930	0.2372	1.0027	0.0759

The underlined values in this table are obtained from the true model, i.e., the S-SMN-RMEM

Table 4 Performance of estimators under C-SMN-RMEM datasets

n	SMN type	$\phi_\varepsilon/\phi_\delta = 1$						$\phi_\varepsilon/\phi_\delta = 0.2$					
		λ		α		β		λ		α		β	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
20	N-SMN	3.0074	0.3089	1.8437	1.1303	1.0486	0.3669	3.0113	0.3124	1.9413	0.6844	1.0184	0.2172
	T-SMN	3.0059	0.2915	1.8808	0.9552	1.0362	0.3069	3.0082	0.2915	1.9623	0.5756	1.0126	0.1798
	S-SMN	3.0059	0.2851	1.8764	0.9678	1.0379	0.3118	3.0083	0.2853	1.9600	0.5833	1.0128	0.1833
	C-SMN	3.0050	0.2833	1.8809	0.9503	1.0362	0.3051	3.0058	0.2830	1.9568	0.5737	1.0131	0.1801
50	N-SMN	2.9949	0.1880	1.9306	0.6449	1.0214	0.2084	2.9990	0.1980	1.9660	0.4115	1.0109	0.1307
	T-SMN	2.9949	0.1768	1.9580	0.5199	1.0121	0.1663	2.9966	0.1851	1.9778	0.3481	1.0069	0.1093
	S-SMN	2.9950	0.1713	1.9570	0.5135	1.0125	0.1644	2.9967	0.1807	1.9780	0.3470	1.0069	0.1091
	C-SMN	2.9950	0.1700	1.9587	0.5070	1.0121	0.1625	2.9966	0.1781	1.9779	0.3428	1.0066	0.1078
100	N-SMN	3.0015	0.1383	1.9632	0.4441	1.0125	0.1438	2.9982	0.1322	1.9831	0.2892	1.0058	0.0923
	T-SMN	2.9990	0.1299	1.9805	0.3699	1.0068	0.1181	3.0026	0.1278	1.9896	0.2379	1.0036	0.0748
	S-SMN	3.0011	0.1269	1.9794	0.3665	1.0072	0.1174	3.0022	0.1223	1.9881	0.2363	1.0039	0.0747
	C-SMN	2.9997	0.1261	1.9808	0.3576	1.0068	0.1145	3.0021	0.1214	1.9898	0.2318	1.0034	0.0735

The underlined values in this table are obtained from the true model, i.e., the C-SMN-RMEM

We first generate 1,000 datasets with sample sizes $n = 50$ and 100 from model (4) under normal distribution. Similar to Vanegas and Cysneiros (2010), we shift the observed value $x_t^{(i)}$ to $x_t^{(i)} + \lambda d$, where $t = n/2$ and $d = 0, 0.5, 1, \ldots, 5$ to guarantee the presence of one outlier in the individuals. For each data set, we calculate all the five estimators (ML estimators based on four type SMN-RMEMs, and the RC estimators) of α, β and λ under the shifted and non-shifted data, respectively. Then, we compute the relative changes of the estimates (i.e., $|(Est_{(s)} - Est)/Est|$, where Est is the estimate under non-shifted data, $Est_{(s)}$ is the estimate under shifted data).

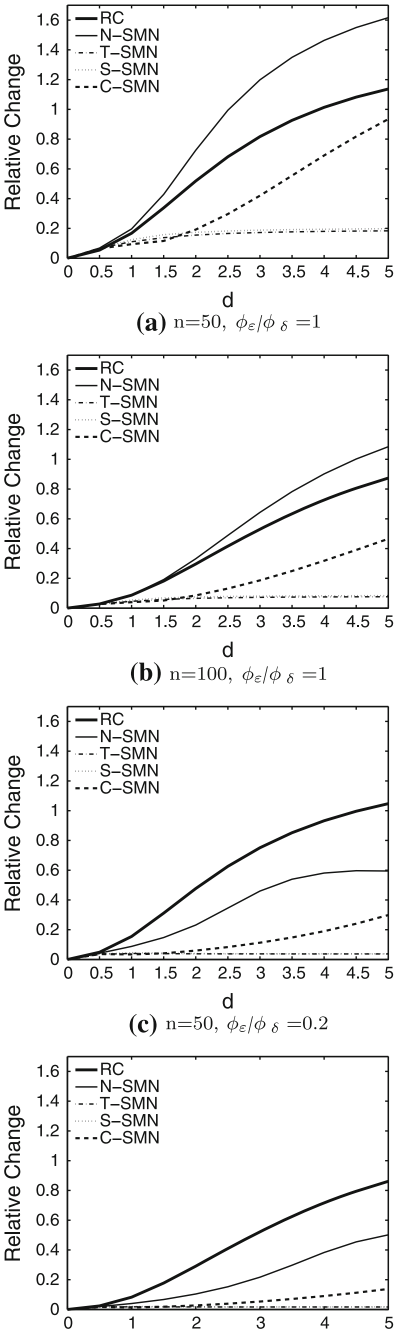
Figures 1, 2 and 3, respectively display the average relative changes on the estimates $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\lambda}$ at different values of d under all the five estimation methods. In all situations, the change ratios of RC and N-SMN estimates increase with d, which indicates that the influence of the outlier become serious when d increases for the RC and N-SMN estimates. On the contrary, the relative changes on the estimates based on T-SMN and S-SMN models are almost not increasing with d. Though the change ratios on the estimates based on C-SMN model show a slightly increasing trend with d, they are still much smaller than those on the RC and N-SMN estimates. As the sample size n increases, we find that the influences of the outlier on the estimates become smaller for all the five estimates. However, the advantage of the robustness based on heavy-tailed models is still obvious. Thus, we draw a conclusion from the simulations that ML estimation method based on heavy-tailed SMSN-RMEM is more appealing since it can present more robustness compared to the traditional normal ones and the RC method.

5 Application

In this section, we consider the CSFII data (Thompson et al. 1992) as a numerical example. This dataset has also been used by Carroll et al. (2006) as an additional information to analyze the NHANES data (Jones et al. 1987). The CSFII data contains the 24-h recall measures, as well as three additional 24-h recall phone interviews of 1,722 women who were recorded about their diet habits. We consider the calorie intake/5,000 as ξ, and the saturated fat intake/100 as η. Instead of ξ and η, the nutrition variables x and y are computed by four 24-h recalls, which are supposed to follow model (4) with $p = q = 4$.

Figure 4 displays the linear tendency between the average calories ($\bar{x}$) and the average saturated fat ($\bar{y}$). The QQ plot of the differences between replicates x_t and between replicates y_t are given in Fig. 5a, b, respectively, which show that non-normality is evident in the presence of heavier-than-normal tails. If the CSFII data we used follows a normal RMEM, then it is true that $\{u_t = (Z_t - \boldsymbol{\mu})^\top \Sigma^{-1}(Z_t - \boldsymbol{\mu}),\ t = 1, \ldots, 1{,}722\}$ are mutually independent and follow a chi-square distribution with 8 degrees of freedom. By applying the Wilson-Hilferty transformation (Johnson et al. 1994), we obtain a set of iid variables $\{r_t = 3u_t^{1/3} - 35/6,\ t = 1, \ldots, 1{,}722\}$ which approximately follows the standard normal distribution. The QQ plot of $\{r_t,\ t = 1, \ldots, 1{,}722\}$ is shown in Fig. 5c, which gives obvious evidence against the normal assumption.

Fig. 1 Average relative changes of $\widehat{\alpha}$ under five estimation methods



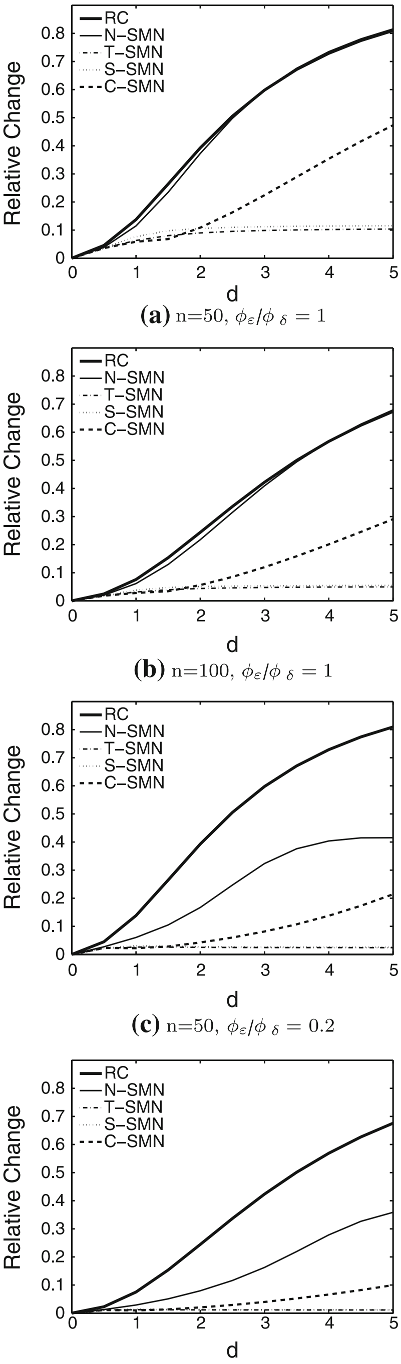
(a) n=50, $\phi_\varepsilon/\phi_\delta$ =1

(b) n=100, $\phi_\varepsilon/\phi_\delta$ =1

(c) n=50, $\phi_\varepsilon/\phi_\delta$ =0.2

(d) n=100, $\phi_\varepsilon/\phi_\delta$ =0.2

Fig. 2 Average relative changes of $\widehat{\beta}$ under five estimation methods



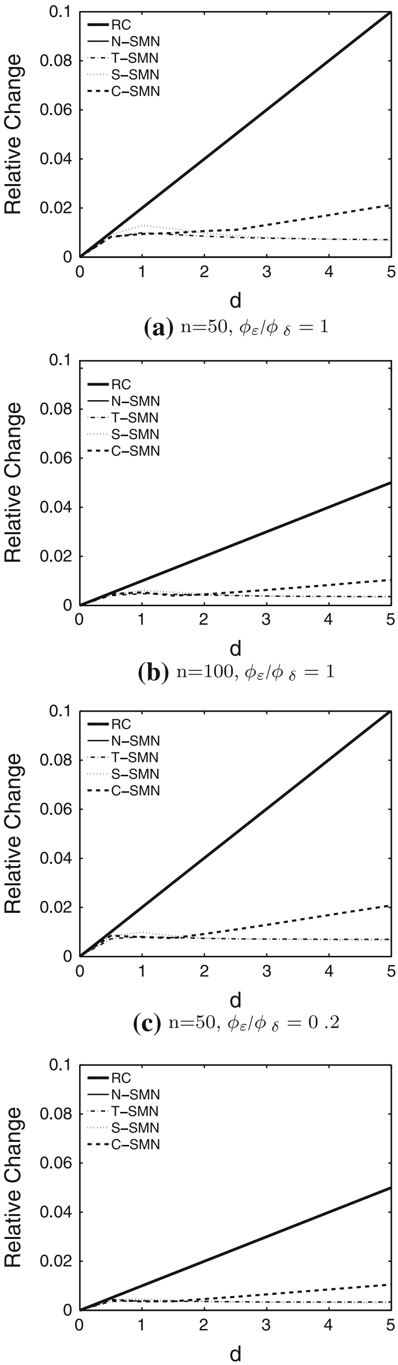
(a) n=50, $\phi_\varepsilon/\phi_\delta = 1$

(b) n=100, $\phi_\varepsilon/\phi_\delta = 1$

(c) n=50, $\phi_\varepsilon/\phi_\delta = 0.2$

(d) n=100, $\phi_\varepsilon/\phi_\delta = 0.2$

Fig. 3 Average relative changes of $\widehat{\lambda}$ under five estimation methods



(a) n=50, $\phi_\varepsilon/\phi_{\ \delta} = 1$

(b) n=100, $\phi_\varepsilon/\phi_{\ \delta} = 1$

(c) n=50, $\phi_\varepsilon/\phi_{\ \delta} = 0\ .2$

(d) n=100, $\phi_\varepsilon/\phi_{\ \delta} = 0.2$

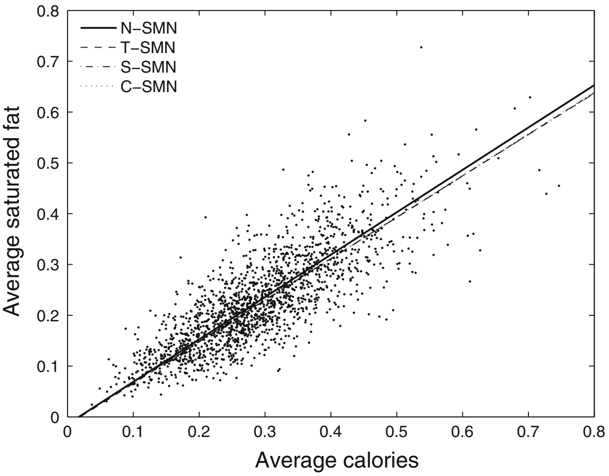


Fig. 4 The linear tendency and four fitted lines between the average calories and the average saturated fat

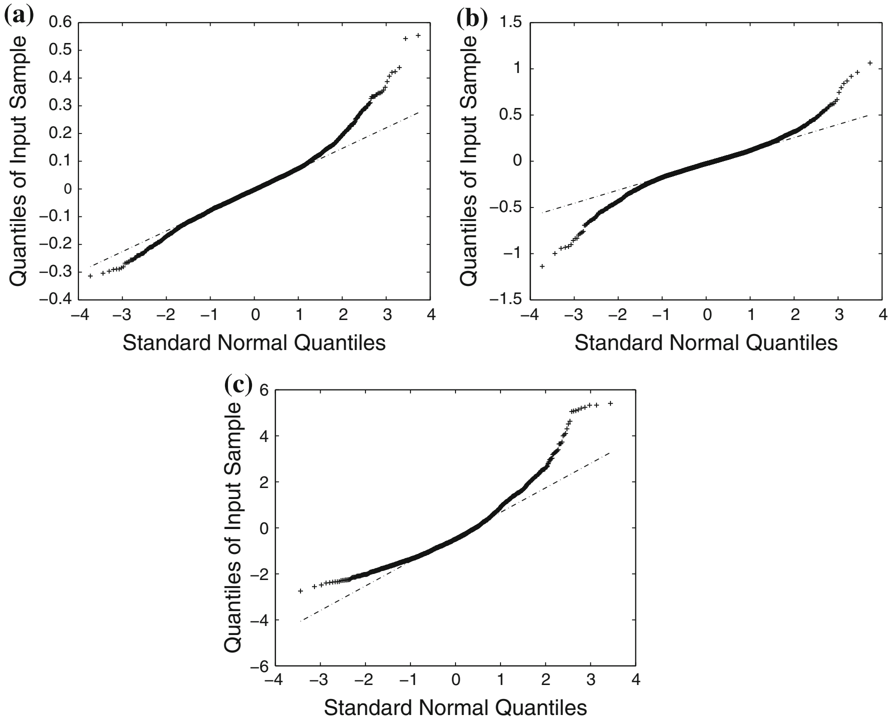


Fig. 5 QQ plots for the CSFII data: **a** differences between replicates of calorie intakes, **b** differences between replicates of saturated fat intakes, and **c** transformation of the Mahalanobis distances

Now we consider the ML estimates for the CSFII data based on RMEM under four proposed SMN distributions. The degrees of freedom for T-SMN, S-SMN, and C-SMN distributions are selected by the Schwarz information criterion (Schwarz

Table 5 Parameter estimators of the CSFII data under SMN-RMEM

Parameter	N-SMN	T-SMN $\nu = 4.3$	S-SMN $\nu = 1.3$	C-SMN $\nu = 0.29, \gamma = 0.22$
λ	0.2897 (0.0025)	0.2677 (0.0023)	0.2701 (0.0023)	0.2706 (0.0023)
α	−0.0143 (0.0058)	−0.0159 (0.0048)	−0.0138 (0.0049)	−0.0143 (0.0049)
β	0.8341 (0.0190)	0.8172 (0.0173)	0.8134 (0.0175)	0.8171 (0.0173)
ϕ_δ	0.0073 (0.0001)	0.0039 (0.0001)	0.0024 (0.0001)	0.0035 (0.0001)
ϕ_ε	0.0137 (0.0002)	0.0076 (0.0002)	0.0046 (0.0001)	0.0067 (0.0001)
ϕ_ξ	0.0089 (0.0004)	0.0066 (0.0003)	0.0039 (0.0002)	0.0055 (0.0002)
AIC	−20,842	−23,182	−23,066	−22,954
BIC	−20,809	−23,144	−23,028	−22,910

The T-SMN-RMEM has the smallest values of BIC and standard errors of location parameters which are underlined in the table

1978). We plot the profile log-likelihood functions for the three models in Fig. 6. By getting the largest values of the profile log-likelihood, the degrees of freedom are found as $\nu = 4.3$ for T-SMN, $\nu = 1.3$ for S-SMN, and $\nu = 0.29$, $\gamma = 0.22$ for C-SMN.

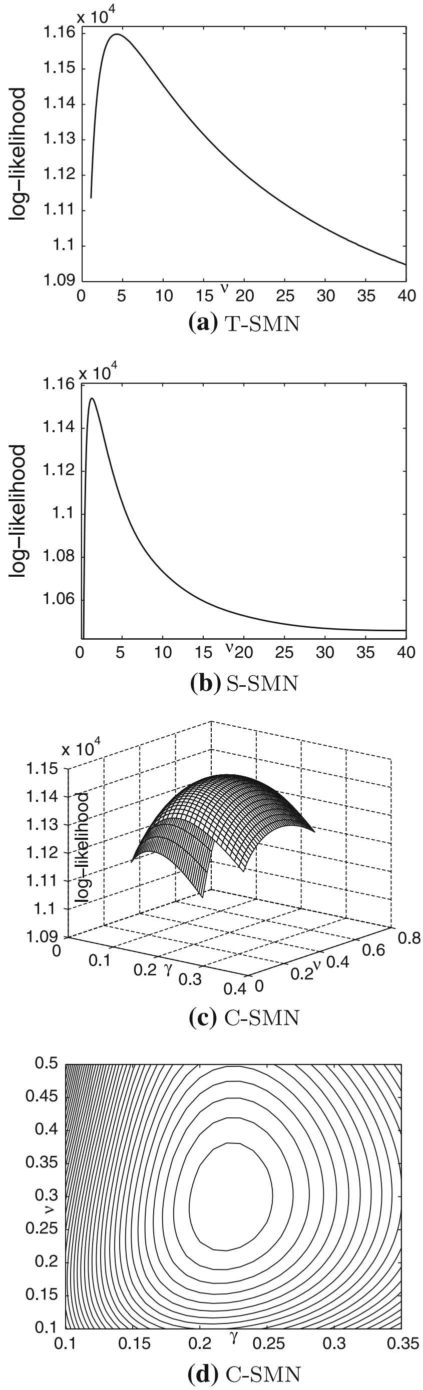
Table 5 gives the ML estimates of parameter $\boldsymbol{\theta}$ with standard errors in parenthesis, and also the AIC (Akaike 1974) and BIC (Schwarz 1978) values of the four SMN-RMEMs. The estimates of the scale parameters are not comparable among different distributions due to the different scales. Note that, the original AIC is used for model selection since the four type distributions we discussed are all members of the SMN distribution class. Compared with the conditional AIC (Vaida and Blanchard 2005), we prefer original AIC in this situation. First, in the RMEM, the most interest is in the population parameters α and β, and not in the individual clusters. Second, as we mentioned, the SMN distributions belong to the elliptical distribution family. The reason we use the SMN form to represent Student-t, slash and contaminated normal distribution is its hierarchical structure makes the EM algorithm become feasible. However, instead of the hierarchical form, the elliptical structure is still the major form of the RMEM model when we do other statistical inference.

From Table 5, we find that the standard errors of λ, α and β, and the values of AIC and BIC under the three heavy-tailed distributions are always smaller than those under the normal one, which indicates that the heavy-than-normal RMEMs fit the data better than N-SMN-RMEM. Moreover, it is suggested that T-SMN-RMEM is the best one among the four models, since it has the smallest values of BIC and standard errors of location parameters. It should be noted that the estimates of β under three heavy-tailed models are all smaller than that under the normal one. This attenuation phenomenon is displayed in Fig.4, in which four regression lines between the average calories and the average saturated fat are plotted, based on the four models, respectively.

6 Conclusions

In this work, we have discussed the ML estimations of the proposed SMN-RMEM. A major advantage of SMN model is its flexibility, due to it contains different types

Fig. 6 Degrees of freedom versus the profile log-likelihood under three heavy-tailed RMEMs, **a** T-SMN, **b** S-SMN, **c, d** C-SMN



(a) T-SMN

(b) S-SMN

(c) C-SMN

(d) C-SMN

of distributions, which offers us the opportunity to compare with each other. Iterative equations are obtained to estimate the parameters of the model by the EM algorithm method. It is important to emphasize the capacity of this model to attenuate outlying observations by using heavy-tailed SMN distributions. Monte Carlo simulations displayed the robustness of heavy-tailed SMN-RMEM. A real data analysis also confirms some robustness aspects of our SMN-RMEM.

Acknowledgments This research was supported by National Science Foundation of China (Grant No. 11171065) and Natural Science Foundation of Jiangsu Province of China (Grant No. BK2011058). We are very grateful to the editor and reviewers for their helpful comments and suggestions which largely improve our work.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723

Andrews DF, Mallows CL (1974) Scale mixtures of normal distributions. J R Stat Soc B 36:99–102

Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM (2006) Measurement error in nonlinear models: a modern perspective, 2nd edn. Chapman and Hall, Boca Raton

Chan LK, Mak TK (1979) Maximum likelihood estimation of a linear structural relationship with replication. J R Stat Soc B 41:263–268

Cheng CL, Van Ness JW (1999) Statistical regression with measurement error. Arnold, London

Cornish EA (1954) The multivariate *t* distribution associated with a set of normal standard deviates. Aust J Phys 7:531–542

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc B 39:1–38 (with discussion)

Dunnett CW, Sobel M (1954) A bivariate generalization of Student's *t* distribution with tables for certain cases. Biometrika 41:153–169

Fang KT, Kotz S, Ng KW (1990) Symmetrical multivariate and related distributions. Chapman and Hall, London

Fuller WA (1987) Measurement error models. Wiley, New York

Harville DA (1997) Matrix algebra from a statistician's perspective. Springer, New York, pp 98–101

Isogawa Y (1985) Estimating a multivariate linear structural relationship with replication. J R Stat Soc B 47:211–215

Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions, 2nd edn. Wiley, New York

Jones DY, Schatzkin A, Green SB, Block G, Brinton LA, Ziegler RG, Hoover R, Taylor PR (1987) Dietary fat and breast cancer in the National Health and Nutrition Examination Survey I: epidemiologic follow-up study. J Natl Cancer Inst 79:465–471

Lachos VH, Angolini T, Abanto-Valle CA (2011) On estimation and local influence analysis for measurement errors models under heavy-tailed distributions. Stat Papers 52:567–590

Lachos VH, Labra FV, Bolfarine H, Ghosh P (2010) Multivariate measurement error models based on scale mixtures of the skew-normal distribution. Statistics 44:541–556

Lange KL, Sinsheimer JS (1993) Normal/independent distributions and their applications in robust regression. J Comput Graph Stat 2:175–198

Lin N, Bailey BA, He XM, Buttlar WG (2004) Adjustment of measuring devices with linear models. Technometrics 46:127–134

McLachlan GL, Krishnan T (1997) The EM algorithm and extensions. Wiley, New York

Osorio F, Paula GA, Galea M (2009) On estimation and influence diagnostics for the Grubb's model under heavy-tailed distributions. Comput Stat Data Anal 53:1249–1263

Pinheiro JC, Liu C, Wu YN (2001) Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate *t* distribution. J Comput Graph Stat 10:249–276

Reiersol O (1950) Identifiability of a linear relation between variables which are subject to errors. Econometrica 18:375–389

Rogers WH, Tukey JW (1972) Understanding some long-tailed symmetrical distributions. Stat Neerlandica 26:211–226

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Thompson FE, Sowers MF, Frongillo EA, Parpia BJ (1992) Sources of fiber and fat in diets of US women aged 19–50: implications for nutrition education and policy. Am J Public Health 82:695–718

Tukey JW (1960) A survey of sampling from contaminated distributions. In: Olkin I (ed) Contributions to probability and statistics. Standford University Press, Stanford, pp 448–485

Vaida F, Blanchard S (2005) Conditional akaike information for mixed-effect models. Biometrika 92:321–370

Vanegas LH, Cysneiros FJA (2010) Assesment of diagnostic procedures in symmetrical nonlinear regression models. Comput Stat Data Anal 54:1002–1016