Sigma J Eng & Nat Sci 36 (2), 2018, 523-538

Sigma Journal of Engineering and Natural Sciences
Sigma Mühendislik ve Fen Bilimleri Dergisi

Research Article

ESTIMATING THE MISSING VALUE IN ONE-WAY ANOVA UNDER LONG-TAILED SYMMETRIC ERROR DISTRIBUTIONS

Demet AYDIN*[1], Birdal ŞENOĞLU[2]

[1]*Sinop University, Department of Statistics, SINOP;* ORCID:0000-0002-3491-8392
[2]*Ankara University, Department of Statistics, ANKARA;* ORCID:0000-0003-3707-2393

Received: 15.06.2017 Revised: 19.09.2017 Accepted: 13.02.2018

ABSTRACT

In practice, missing values are widely seen and create serious problems in almost all statistical analysis. In this study, to deal with missing values, we propose estimators for missing value in one-way analysis of variance (ANOVA) when the distribution of error terms is long-tailed symmetric (LTS). We use methodologies known as maximum likelihood (ML), modified maximum likelihood (MML) and least squares (LS) in estimating missing value. Expectation and maximization (EM) algorithm is used for computing ML estimate of missing value. We compare the efficiencies of LS, ML and MML estimators of missing value via Monte Carlo simulation study. Simulation results show that ML estimator of missing value is the most efficient among the others. The usefulness of the proposed estimators is illustrated by peak discharge data example taken from civil engineering.

Keywords: Missing value, one-way ANOVA, LTS distribution, EM algorithm, MML methodology.

1. INTRODUCTION

In real life problems, certain observations are sometimes missing due to the various reasons. For example, the experimenter fails to record some data, crops are destroyed in some plots, a patient withdraws from the treatment, and one or more animals die in the course of the experiment. Therefore, the problem of missing value is quite common in almost all type of research. In the presence of missing values, their omission naturally affects the method of analysis, so they should be estimated before analyzing the data [9, 11].

Estimation of the missing value was first studied by Allan and Whishart [1]. They obtained the estimate of the missing value by minimizing the error sum of squares (*SSE*) in randomized block design and Latin square design. Their method was extended by Yates [45] to the several missing values case. They use iterative methods for estimating the missing values. Bartlett [5] proposed a new method to obtain the *LS* estimates of the missing values by using the non-iterative method called as analysis of covariance (*ANCOVA*). A general non-iterative method was proposed by Hartley [16] for estimating the missing value. Healy and Westmacott [17] proposed another method based on iterative techniques for estimating more than one missing values. It should be noted that *LS* method based on the idea of minimizing *SSE* with respect to the unknown

* Corresponding Author: e-mail: daydin@sinop.edu.tr, tel: (368) 271 55 16

parameters and to the missing values was used in these studies. In the estimation procedure, the missing values are treated as unknown parameters. Little and Rubin [23] estimated the location and the scale pa rameters of the normal distribution in the presence of missing value. They obtained *ML* estimates of the unknown parameters by using the approach where the missing value is treated as a parameter, see also [5, 2, 44, 30, 33, 20, 37, 38, 32, 3]. It can be seen from these studies that this approach is computationally feasible for many missing data problems.

In this study, we obtain the estimators of the missing value in one-way *ANOVA* by using the *ML* and the *MML* methodologies when the distribution of the error terms is *LTS*. In the estimation procedure, similar to the earlier studies, we use the approach where the missing value is treated as parameter.

In the literature, there are various missing data mechanisms that describe how the probability of a missing value relates to the data. Commonly used ones are missing completely at random (*MCAR*), missing at random (*MAR*) and missing not at random (*MNAR*), see [13, 34]. The missingness mechanism does not depend on the observed or the unobserved data in *MCAR*. Missingness depends only on the observed data in *MAR*. The missing data mechanism is *MNAR* when neither *MAR* nor *MCAR* hold. In this study, we use the missingness mechanism known as *MCAR*, since the reason for the occurrence of missing value in one-way *ANOVA* does not depend on any observed or unobserved data.

The remaining of the paper is organized as follows. In Section 2, we introduce one-way *ANOVA* model, *LTS* distribution and their properties. We obtain the *ML* estimate equation of the missing value in Section 3. Section 4 gives *MML* estimator of the missing value. The results of Monte-Carlo simulation study are presented in Section 5. A real data set taken from the literature is analyzed to illustrate the performance of the proposed estimators in Section 6. We give conclusions at the end of the paper.

2. THE STATISTICAL MODEL

Consider the following one-way *ANOVA* model

$$y_{ij} = \mu_i + \varepsilon_{ij},\ i = 1,2,\cdots,a;\ j = 1,2,\cdots,n \quad (1)$$

where y_{ij} is the *j*th observation in the *i*th treatment, μ_i is the mean of the *i*th treatment, *a* is the number of treatment levels, *n* is the number of observation in each treatment level and ε_{ij} are independently and identically distributed (*iid*) random error terms. Here, we assume that the model is a fixed-effects model.

In the context of *ANOVA*, traditionally, error terms are assumed to be normally distributed. However, non-normal distributions are encountered more frequently in practice; see for example [28, 15, 39, 12, 19, 42]. When the distribution of the error terms is non-normal, the *LS* estimators lose their efficiency very quickly. Unfortunately, the effect of a violation of the normality assumption on the efficiencies of the estimators is frequently overlooked. One-way of handling non-normal data is to use Box and Cox [8] normalizing transformation. However, all non-normal data cannot be amenable to this transformation. Also, it is often difficult to interpret the transformed parameters, see Bickel and Doksum [6]. In this study, we use original data rather than the transformed data.

Different than the earlier studies, we assume that the error terms are *iid* and have the family of distributions represented by *LTS*. The probability density function (*pdf*) of the *LTS* distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{q}B\left(\frac{1}{2},\ p-\frac{1}{2}\right)}\left(1 + \frac{(x-\mu)^2}{q\sigma^2}\right)^{-p},\ -\infty < x < \infty,\ \mu \in \mathbb{R},\ \sigma > 0 \quad (2)$$

where p is the shape para meter, $q = 2p - 3$ for p≥ 2, see Tiku and Kumra [41]. μ is the location parameter, σ is the scale parameter and $B(\cdot,\cdot)$ is the beta function. Mean, variance and kurtosis (β_2) of the *LTS* distribution are given below

$$E(X) = \mu, Var(X) = \sigma^2 \text{ and } \beta_2 = \frac{3(p-3/2)}{(p-5/2)}. \tag{3}$$

In Table 1, kurtosis values of the *LTS* distribution are also presented for better evaluating the shape of the distribution for different p values.

Table 1. The kurtosis values for the *LTS* distribution for certain selected values of p

p:	2.5	3.0	3.5	5.0	10	∞
β_2:	∞	9.0	6.0	4.2	3.4	3.0

Note that $T = \sqrt{\nu/q}\frac{(X-\mu)}{\sigma}$ has the Student-*t* distribution with degrees of freedom $\nu = 2p - 1$. The shape parameter p is assumed to be known throughout the estimation process.

One of the most frequently used non-normal symmetric distributions is *LTS*, see [42, 43, 21, 27, 35, 36]. It is commonly used in modelling the data sets having long tails and outliers. Here, it should be noted that outliers exist in the direction of the long tails.

Plots of the *LTS* distribution for some representative values of the shape parameter p and the standard normal distribution are given in Figure 1. In the plots, μ and σ were taken to be 0 and 1, respectively.

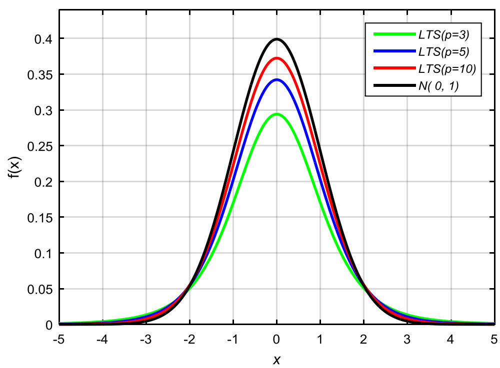

Figure 1. Plots of the *LTS* distribution for the different values of the shape parameter p and standard normal distribution

It is clear from Table 1 and Figure 1 that the *LTS* distribution is plausible alternative to the normal distribution. The kurtosis is always greater than 3; however, *LTS* distribution reduces to the well-known normal distribution as the shape parameter p goes to infinity.

3. THE *ML* ESTIMATOR OF MISSING VALUE

Suppose that the observation y_{kl} which is the *l*th observation in *k*th treatment is missing. The likelihood function (L) can be factorized as a product of two components. The first component is for the observed values (y_{ij}; $i \neq k$, $j \neq l$) and the second component is for the missing value

(y_{kl}), see Rao and Toutenburg [31]. Then, log-likelihood function $(\ln L)$ is written as shown below when the distribution of the error terms is *LTS*.

$$\ln L \propto -N\ln\sigma - p\sum_{\substack{i=1\\ i\neq k}}^{a}\sum_{j=1}^{n}\ln\left(1+\frac{(y_{ij}-\mu_i)^2}{q\sigma^2}\right) - p\sum_{\substack{j=1\\ j\neq l}}^{n}\ln\left(1+\frac{(y_{kj}-\mu_k)^2}{q\sigma^2}\right)$$
$$-p\ln\left(1+\frac{(m-\mu_k)^2}{q\sigma^2}\right) \tag{4}$$

The *ML* estimator of the model parameters (μ_i, σ) and the missing value (m) are the solutions of the following likelihood equations

$$\frac{\partial \ln L}{\partial \mu_i} = \frac{2p}{q\sigma}\sum_{j=1}^{n} g(z_{ij}) = 0 \tag{5}$$

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma}\sum_{\substack{j=1\\ j\neq l}}^{n} g(z_{kj}) + \frac{2p}{q\sigma} g(z_{kl}) = 0 \tag{6}$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{N}{\sigma} - \frac{2p}{q\sigma}\sum_{\substack{i=1\\ i\neq k}}^{a}\sum_{j=1}^{n} z_{ij} g(z_{ij}) - \frac{2p}{q\sigma}\sum_{\substack{j=1\\ j\neq l}}^{n} z_{kj} g(z_{kj}) - \frac{2p}{q\sigma} z_{kl} g(z_{kl}) = 0 \tag{7}$$

$$\frac{\partial \ln L}{\partial m} = \frac{2p}{q\sigma} g(z_{kl}) = 0\,. \tag{8}$$

Here, $g(z_{ij}) = \frac{z_{ij}}{\left(1+\frac{1}{q}z_{ij}^2\right)}$ and $z_{ij} = \frac{(y_{ij}-\mu_i)}{\sigma}$. These equations do not have explicit solutions because of the nonlinear functions $g(.)$. Therefore, we resort to iterative methods, such as Newton–Raphson, iteratively reweighting algorithm (*IRA*) or *EM* algorithm. In this study, *EM* algorithm is used for computing the *ML* estimate of the missing value. The *EM* algorithm has reliable global convergence under fairly general conditions; see McLachlan and Krishnan [25] for more detailed information. The *EM* algorithm introduced by Dempster et al. [10] is a very popular iterative method for incomplete data problems. It has two steps called as the expectation (*E*) and the maximization (*M*). Step *E* is defined as the conditional expectation of the lnL for complete data given observed incomplete data and the current value of the parameters. Step *M* is defined as the maximization of the conditional expectation obtained in step *E* with respect to the unknown parameters. *E* and *M* steps are repeated until the following condition is satisfied,

$$\left|m_{ML}^{(t)} - m_{ML}^{(t-1)}\right| \le \varepsilon,\ t = 1,2,\cdots,. \text{ and } (\varepsilon > 0). \tag{9}$$

Here, $m_{ML}^{(t)}$ satisfying this condition is the *ML* estimate of the missing value *m*. To solve the equations in (5)-(8) by using the *EM* algorithm, we define the complete data vector $\boldsymbol{X}$ as $\boldsymbol{X} = (\boldsymbol{Y},\boldsymbol{U})$. Here, $\boldsymbol{Y} = (\boldsymbol{Y_1},\boldsymbol{Y_2},\cdots,\boldsymbol{Y_a})$ is incomplete data vector, $\boldsymbol{Y_i} = (Y_{i1},Y_{i2},\cdots,Y_{in})$ is the vector of observations in *i*th treatment $(i \neq k)$, $\boldsymbol{Y_k} = (Y_{k1},Y_{k2},\cdots,Y_{kl-1},m,Y_{kl+1},\cdots,Y_{kn})$ is the vector of observations in *k*th treatment including missing value, $\boldsymbol{U} = (\boldsymbol{U_1},\boldsymbol{U_2},\cdots,\boldsymbol{U_a})$ is the missing variable vector that would never be observable for complete data vector $\boldsymbol{X}$ and $\boldsymbol{U_i} = (U_{i1},U_{i2},\cdots,U_{in})$ is the vector of missing variables corresponding to the *i*th treatment. Here, it should be realized that *m* is the missing observation while $\boldsymbol{U}$ is the vector of missing (or latent) variables, therefore it represents an artificial vector of additional missing variables.

The joint distribution of $\boldsymbol{Y}$ and $\boldsymbol{U}$ can be written as the product of the marginal distribution of $\boldsymbol{U}$ and the conditional distribution of $\boldsymbol{Y}$ given $\boldsymbol{U}$ as shown below

$$f(y_{ij},u_{ij}) = f(y_{ij}|u_{ij})f(y_{ij}). \tag{10}$$

It is obvious that assuming non-normal distributions for both $\boldsymbol{Y}$ and $\boldsymbol{U}$ provides great flexibility for modeling the bivariate data. Here, we assume that *pdf* of Y_{ij} is *LTS* with parameters μ_i and σ and *pdf* of U_{ij} is χ_ν^2 with degrees of freedom $\nu = 2p-1$. Then the conditional density function of Y_{ij} given U_{ij} is obtained to be $N(\mu_i, q\sigma^2/u_{ij})$. In the literature, there exists studies assuming the distribution of $\boldsymbol{Y}$ is Student-t, see for example [22, 24, 29]. Then, we define the likelihood function for the complete data with one missing value as shown below

$$L_c(\boldsymbol{\theta}) = L_{y|u}(\boldsymbol{\theta})L_u \tag{11}$$

where

$$L_{y|u}(\boldsymbol{\theta}) = \prod_{i=1}^{a}\prod_{j=1}^{n}\frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}}exp\left(-\frac{1}{2}\frac{(y_{ij}-\mu_i)^2}{q\sigma^2/u_{ij}}\right)$$
$$= \prod_{\substack{i=1\\ i\neq k}}^{a}\prod_{j=1}^{n}\frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}}exp\left(-\frac{1}{2}\frac{(y_{ij}-\mu_i)^2}{q\sigma^2/u_{ij}}\right)$$
$$\cdot\prod_{\substack{j=1\\ j\neq l}}^{n}\frac{1}{\sqrt{2\pi(q\sigma^2)/u_{kj}}}exp\left(-\frac{1}{2}\frac{(y_{kj}-\mu_k)^2}{q\sigma^2/u_{kj}}\right)$$
$$\cdot\frac{1}{\sqrt{2\pi(q\sigma^2)/u_{kl}}}exp\left(-\frac{1}{2}\frac{(m-\mu_k)^2}{q\sigma^2/u_{kl}}\right)$$

and

$$L_u = \prod_{i=1}^{a}\prod_{j=1}^{n}\frac{1}{\Gamma(\nu/2)2^{\nu/2}}u_{ij}^{(\nu-2)/2}\,exp\left(-\frac{1}{2}u_{ij}\right)$$
$$= \prod_{\substack{i=1\\ i\neq k}}^{a}\prod_{j=1}^{n}\frac{1}{\Gamma(\nu/2)2^{\nu/2}}u_{ij}^{(\nu-2)/2}\,exp\left(-\frac{1}{2}u_{ij}\right)$$
$$.\prod_{\substack{j=1\\ j\neq l}}^{n}\frac{1}{\Gamma(\nu/2)2^{\nu/2}}u_{kj}^{(\nu-2)/2}exp\left(-\frac{1}{2}u_{kj}\right)$$
$$\cdot\frac{1}{\Gamma(\nu/2)2^{\nu/2}}u_{kl}^{(\nu-2)/2}exp\left(-\frac{1}{2}u_{kl}\right).$$

Here, $\boldsymbol{\theta} = (\mu_i, \mu_k, \sigma, m)$ is the vector of unknown parameters and *m*. To obtain the solutions of the likelihood equations in (5)-(8) with respect to the unknown parameters and *m*, we apply the *EM* algorithm whose steps are given below.

Step E: Conditional expectation of the $lnL_c(\boldsymbol{\theta})$ given $\boldsymbol{y} = (y_{11}, y_{12}, \cdots, y_{1n}; \cdots; y_{a1}, y_{a2}, \cdots, y_{an})$ is obtained as follows

$$\Phi\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(t-1)}\right) = E_{\boldsymbol{\theta}^{(t-1)}}(lnL_c(\boldsymbol{\theta})\,|\boldsymbol{y})$$
$$= -Nln\left(\Gamma(\nu/2)2^{\nu/2}\right) - \frac{(a-1)n}{2}ln\left(q\sigma^2/u_{ij}\right) - \frac{(n-1)}{2}ln\left(q\sigma^2/u_{kj}\right)$$
$$-\frac{1}{2}ln(q\sigma^2/u_{kl}) - \frac{1}{2q\sigma^2}\sum_{\substack{i=1\\ i\neq k}}^{a}\sum_{j=1}^{n}(y_{ij}-\mu_i)^2E_{\boldsymbol{\theta}^{(t-1)}}\left(u_{ij}\middle|y_{ij}\right)$$
$$-\frac{1}{2q\sigma^2}\sum_{\substack{j=1\\ j\neq l}}^{n}(y_{kj}-\mu_k)^2E_{\boldsymbol{\theta}^{(t-1)}}\left(u_{kj}\middle|y_{kj}\right)$$
$$-\frac{1}{2q\sigma^2}(m-\mu_k)^2E_{\boldsymbol{\theta}^{(t-1)}}(u_{kl}|m)$$
$$-\frac{1}{2}\left(\sum_{\substack{i=1\\ i\neq k}}^{a}\sum_{j=1}^{n}E_{\boldsymbol{\theta}^{(t-1)}}\left(u_{ij}\middle|y_{ij}\right) + \sum_{\substack{j=1\\ j\neq l}}^{n}E_{\boldsymbol{\theta}^{(t-1)}}\left(u_{kj}\middle|y_{kj}\right) + E_{\boldsymbol{\theta}^{(t-1)}}(u_{kl}|m)\right)$$
$$+\frac{\nu-2}{2}\left(\sum_{\substack{i=1\\ i\neq k}}^{a}\sum_{j=1}^{n}E_{\boldsymbol{\theta}^{(t-1)}}\left(lnu_{ij}\middle|y_{ij}\right) + \sum_{\substack{j=1\\ j\neq l}}^{n}E_{\boldsymbol{\theta}^{(t-1)}}\left(lnu_{kj}\middle|y_{kj}\right) + E_{\boldsymbol{\theta}^{(t-1)}}(lnu_{kl}|m)\right) \tag{12}$$

Step M: We take the derivatives of Φ with respect to the unknown parameters and the missing value *m* and equate them to zero. Then we obtain the following likelihood equations

$$\frac{\partial\Phi\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(t-1)}\right)}{\partial\mu_i} = \frac{1}{q\sigma^2}\sum_{j=1}^{n}(y_{ij}-\mu_i)E_{\mu_i^{(t-1)}}\left(u_{ij}\middle|y_{ij}\right) = 0 \tag{13}$$

$$\frac{\partial\Phi\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(t-1)}\right)}{\partial\mu_k} = \frac{1}{q\sigma^2}\sum_{\substack{j=1\\ j\neq l}}^{n}(y_{kj}-\mu_k)E_{\mu_k^{(t-1)}}\left(u_{kj}\middle|y_{kj}\right)$$

$$+\frac{1}{q\sigma^2}(m-\mu_k)E_{\mu_k^{(t-1)}}(u_{kl}|m)=0 \tag{14}$$

$$\frac{\partial\Phi(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t-1)})}{\partial\sigma^2}=-\frac{1}{q\sigma^2}\left(N-\sum_{\substack{i=1\\i\neq k}}^{a}\sum_{j=1}^{n}\frac{\left(y_{ij}-\mu_i\right)^2}{\sigma^2}E_{\mu_i^{(t-1)}}\left(u_{ij}\middle|y_{ij}\right)\right)$$
$$+\frac{1}{q\sigma^2}\sum_{\substack{j=1\\j\neq l}}^{n}\frac{\left(y_{kj}-\mu_k\right)^2}{\sigma^2}E_{\mu_k^{(t-1)}}\left(u_{kj}\middle|y_{kj}\right)$$
$$+\frac{1}{q\sigma^2}\frac{(m-\mu_k)^2}{\sigma^2}E_{\mu_k^{(t-1)}}(u_{kl}|m)=0 \tag{15}$$

$$\frac{\partial\Phi(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t-1)})}{\partial m}=\frac{1}{q\sigma^2}(m-\mu_k)E_{\mu_k^{(t-1)}}(u_{kl}|m)=0. \tag{16}$$

Now, we need the conditional *pdf* of u_{ij} given y_{ij} to obtain the conditional expectations given in the equations (13)-(16). Conditional *pdf* of u_{ij} given y_{ij} can easily be obtained as shown below

$$f\left(u_{ij}\middle|y_{ij}\right)=\frac{\left(1+\frac{\left(y_{ij}-\mu_i\right)^2}{q\sigma^2}\right)^{(\nu+1)/2}}{\Gamma((\nu+1)/2)2^{(\nu+1)/2}}u_{ij}^{(\nu-1)/2}exp\left(-u_{ij}\frac{1}{2}\left(1+\frac{\left(y_{ij}-\mu_i\right)^2}{q\sigma^2}\right)\right). \tag{17}$$

It is clear from (17) that distribution of $u_{ij}|y_{ij}$ is Gamma with $\alpha=(\nu+1)/2$ and $\beta=\frac{1}{2}\left(1+\frac{\left(y_{ij}-\mu_i\right)^2}{q\sigma^2}\right)$ then the conditional expectations for $u_{ij}|y_{ij}$ and $u_{kj}|y_{kj}$ are

$$E_{\mu_i^{(t-1)}}\left(u_{ij}\middle|y_{ij}\right)=\varpi_{ij}^{(t-1)} \text{ and } E_{\mu_k^{(t-1)}}\left(u_{kj}\middle|y_{kj}\right)=\varpi_{kj}^{(t-1)}, \tag{18}$$

where

$$\varpi_{ij}^{(t-1)}=\frac{\nu+1}{1+\frac{1}{q}\delta_{ij}^2},\ \varpi_{kj}^{(t-1)}=\frac{\nu+1}{1+\frac{1}{q}\delta_{kj}^2},\ \delta_{ij}=\frac{y_{ij}-\mu_i^{(t-1)}}{\sigma^{(t-1)}} \text{ and } \delta_{kj}=\frac{y_{kj}-\mu_k^{(t-1)}}{\sigma^{(t-1)}}.$$

It can easily be seen from equation (16) that $(m-\mu_k)$ is equal to zero, since $u_{kj}|y_{kj}$ is distributed as Gamma and this implies that $E_{\mu_k^{(t-1)}}(u_{kl}|m)>0$. By inserting $(m-\mu_k)=0$ in equations (13)-(15), then we obtain the following equalities

$$\mu_{i,ML}^{(t)}=\frac{\sum_{j=1}^{n}y_{ij}E_{\mu_{i,ML}^{(t-1)}}(u_{ij}|y_{ij})}{\sum_{j=1}^{n}E_{\mu_{i,ML}^{(t-1)}}(u_{ij}|y_{ij})},\ (i=1,2,\cdots,a;i\neq k) \tag{19}$$

$$\mu_{k,ML}^{(t)}=\frac{\sum_{\substack{j=1\\j\neq l}}^{n}y_{kj}E_{\mu_{k,ML}^{(t-1)}}(u_{kj}|y_{kj})}{\sum_{\substack{j=1\\j\neq l}}^{n}E_{\mu_{k,ML}^{(t-1)}}(u_{kj}|y_{kj})}, \tag{20}$$

$$\sigma_{ML}^{(t)}=\sqrt{\frac{\sum_{\substack{i=1\\i\neq k}}^{a}\sum_{j=1}^{n}\left(y_{ij}-\mu_{i,ML}^{(t)}\right)^2E_{\mu_{i,ML}^{(t-1)}}(u_{ij}|y_{ij})+\sum_{\substack{j=1\\j\neq l}}^{n}\left(y_{kj}-\mu_{k,ML}^{(t)}\right)^2E_{\mu_{k,ML}^{(t-1)}}(u_{kj}|y_{kj})}{N-a-1}}. \tag{21}$$

Note that, the dominator of σ is replaced by $N-a-1$ for bias correction. Substituting (18) into (19)-(21), we obtain the following *ML* estimates at the *t*th iteration $(t=1,2,\cdots,)$

$$\mu_{i,ML}^{(t)}=\frac{1}{\omega^{(t-1)}}\sum_{j=1}^{n}w_{ij}^{(t-1)}y_{ij} \tag{22}$$

$$\mu_{k,ML}^{(t)}=\frac{1}{\omega_m^{(t-1)}}\sum_{\substack{j=1\\j\neq l}}^{n}w_{kj}^{(t-1)}y_{kj} \tag{23}$$

$$\sigma_{ML}^{(t)} = \sqrt{\frac{\sum_{\substack{i=1\\i\neq k}}^{a}\sum_{j=1}^{n} w_{ij}^{(t-1)}\left(y_{ij}-\mu_{i,ML}^{(t)}\right)^2 + \sum_{\substack{j=1\\j\neq l}}^{n} w_{kj}^{(t-1)}\left(y_{kj}-\mu_{k,ML}^{(t)}\right)^2}{N-a-1}} \tag{24}$$

where

$$w_{ij}^{(t-1)} = \frac{\varpi_{ij}^{(t-1)}}{q},\ w_{kj}^{(t-1)} = \frac{\varpi_{kj}^{(t-1)}}{q},\ \omega^{(t-1)} = \sum_{j=1}^{n} w_{ij}^{(t-1)} \text{ and } \omega_m^{(t-1)} = \sum_{\substack{j=1\\j\neq l}}^{n} w_{kj}^{(t-1)}.$$

As shown before $(m-\mu_k)=0$, then the *ML* estimate of *m* at the *t*th iteration is given by

$$m_{ML}^{(t)} = \frac{1}{\omega_m^{(t-1)}}\sum_{\substack{j=1\\j\neq l}}^{n} w_{kj}^{(t-1)} y_{kj},\ (t=1,2,\cdots,). \tag{25}$$

Estimates in (22)-(25) are called as *M* estimates with $\rho = -lnf(x)$, see Huber [18]. It is clear that the weight function w_{ij} decreases as δ_{ij} increases. w_{ij} are nonnegative and this depletes the dominant effect of the outliers by giving small weights to them in the direction of the long tails.

To avoid the difficulties encountered in the iterative methods (such as initial value problem, etc.), we can use the methodology known as *MML* proposed by Tiku [40]. In the next section, we give the description of the *MML* methodology and obtain the explicit estimator of the missing value.

4. THE *MML* ESTIMATOR OF MISSING VALUE

MML methodology is used to obtain the explicit estimators of the model parameters by linearizing the nonlinear terms in the likelihood equations.

To obtain the *MML* estimator of *m*, likelihood equations in (6) and (8) are written in terms of the order statistics, since complete sums are invariant to ordering (i.e, $\sum z_i = \sum z_{(i)}$)

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma}\sum_{\substack{j=1\\j\neq l}}^{n} g\left(z_{k(j)}\right) + \frac{2p}{q\sigma} g\left(z_{k(l)}\right) = 0 \tag{26}$$

$$\frac{\partial \ln L}{\partial m} = \frac{2p}{q\sigma} g\left(z_{k(l)}\right) = 0. \tag{27}$$

Incorporating equation (27) into the equation (26), we obtain the likelihood equation for μ_k (mean of the *k*th treatment) given below

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma}\sum_{\substack{j=1\\j\neq l}}^{n} g\left(z_{k(j)}\right) = 0 \tag{28}$$

where $g\left(z_{k(j)}\right) = \frac{z_{k(j)}}{\left(1+\frac{1}{q}z_{k(j)}^2\right)}$ and $z_{k(j)} = \frac{(y_{k(j)}-\mu_k)}{\sigma}$. We then linearize the nonlinear term $g\left(z_{k(j)}\right)$ in (28) by using the first two terms of the Taylor series expansion around the expected values of the order statistics (i.e, $E\left(z_{(i)}\right) = t_{(i)}$)

$$g\left(z_{k(j)}\right) \cong \alpha_j + \beta_j z_{k(j)} \tag{29}$$

where $\alpha_j = \frac{(2/q)t_{(j)}^3}{\left(1+(1/q)t_{(j)}^2\right)^2}$ and $\beta_j = \frac{1-(1/q)t_{(j)}^2}{\left(1+(1/q)t_{(j)}^2\right)^2}$.

Here, it should be noted that approximate $t_{(i)}$ values are obtained from the following equality

$$F\left(t_{(j)}\right) = \frac{1}{\sigma\sqrt{q}B\left(\frac{1}{2},p-\frac{1}{2}\right)}\int_{-\infty}^{t_{(j)}}\left(1+\frac{z^2}{q}\right)^{-p} dz = \frac{j}{n},\ (j=1,2,\cdots,n;\ j\neq l). \tag{30}$$

Substituting (29) into (28), we obtain the following modified likelihood equation

$$\frac{\partial \ln L}{\partial \mu_k} \cong \frac{\partial \ln L^*}{\partial \mu_k} = \frac{2p}{q\sigma}\sum_{\substack{j=1\\ j\neq l}}^{b}\left(\alpha_j + \beta_j z_{k(j)}\right) = 0. \quad (31)$$

The solution of this equation is the following *MML* estimator of the missing value *m* since $(m - \mu_k) = 0$ (see section 3)

$$\hat{m}_{MML} = \frac{1}{w}\sum_{\substack{j=1\\ j\neq l}}^{n}\beta_j y_{k(j)} \quad (32)$$

where $w = \sum_{\substack{j=1\\ j\neq l}}^{n}\beta_j$, see also Aydin and Senoglu [4].

The *MML* estimators have significant advantages; they are asymptotically equivalent to the *ML* estimators and therefore they are fully efficient under the regularity conditions. They are numerically very close in value to the *ML* estimators even for small samples. They are also robust to the outliers and to the non-normality, see Tiku and Akkaya [43].

5. MONTE CARLO SIMULATION STUDY

In this section, we present an extensive Monte Carlo simulation study comparing the performances of the *LS*, the *ML* and the *MML* estimators of the missing value in one-way *ANOVA*. Suppose that y_{kl} (*l*th observation in the *k*th treatment) is missing. It should be remembered that the *LS* estimator of the missing observation *m* is given by

$$\hat{m}_{LS} = \frac{y_{k.}^*}{n-1}$$

where $y_{k.}^*$ is the total of the observations in the *k*th treatment except missing value. We also compare the performances of the estimators of the model parameters $(\hat{\mu}_i,\ \hat{\sigma})$ based on full data with the corresponding performances based on complete data including one imputed missing data to see the effect of imputation on the efficiencies of the estimators.

In the simulations, we take $\mu_i = 0$ $(i = 1,2,\cdots,a)$ and $\sigma = 1$ without loss of generality. We use the transformation $T = \sqrt{v/q}\,\frac{(X-\mu)}{\sigma}$ $(v = 2p - 1)$ to generate the random numbers from the *LTS* distribution. In the simulation study, we use the sample sizes $n = 10$, 15 and 20, and the shape parameters $p = 2$, 2.5, 3, 3.5, 5 and 10. Number of treatments is taken to be 4 just for an illustration.

The *MCAR* mechanism is used to create the missing value. We use methodology proposed by Brand et al. [7] in generating missing value. Missing value can be created in any treatment but we create it in the second one just for an illustration. In this study, simulated values are computed based on $[\![100.000/n]\!]$ Monte Carlo runs. Here, $[\![.]\!]$ represents the integer value function. All the simulations were conducted in MATLAB 2012b.

5.1. Comparison of the Estimators of Missing Value

Three different criteria are used to compare the efficiencies of the estimators of m similar to those given in Engels and Diehr [14]. They are mean deviation (*bias*), mean absolute deviation (*MAD*) and mean-square deviation (*MSD*). See the equalities given below

$$Bias(\hat{m}) = \frac{1}{r}\sum_t^r (m_t - \hat{m}_t),\ MAD(\hat{m}) = \frac{1}{r}\sum_t^r |m_t - \hat{m}_t|$$

and

$$MSD(\hat{m}) = \frac{1}{r}\sum_t^r (m_t - \hat{m}_t)^2 .$$

Here,
$\hat{m}_t$: Estimate value of m in *t*th replication
m_t: The true value of missing value in *t*th replication
r : Number of replication.

The mean deviation is used to assess the *bias*, *MAD* and *MSD* criteria are used to measure the accuracy of the estimators. Simulation results showing the efficiencies of the estimators of missing value are given in Table 2.

Table 2. The *bias*, *MAD* and *MSD* values of *LS*, *ML* and *MML* estimators of *m*

	Bias			*MAD*			*MSD*		
n	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$
				$p = 2$					
10	-0.0144	-0.0134	-0.0137	0.6921	0.6696	0.6724	0.9961	0.9548	0.9582
15	-0.0054	-0.0002	-0.0016	0.6803	0.6598	0.6623	1.0595	1.0237	1.0274
20	0.0370	0.0342	0.0347	0.6694	0.6561	0.6573	1.0271	1.0063	1.0073
				$p = 2.5$					
10	0.0047	0.0039	0.0045	0.7588	0.7460	0.7483	1.1205	1.0929	1.0976
15	-0.0207	-0.0192	-0.0199	0.7236	0.7142	0.7159	0.9898	0.9698	0.9731
20	-0.0092	-0.0081	-0.0085	0.7326	0.7259	0.7267	1.0558	1.0426	1.0446
				$p = 3$					
10	-0.0102	-0.0101	-0.0103	0.7909	0.7820	0.7824	1.1038	1.0808	1.0823
15	0.0128	0.0106	0.0112	0.7601	0.7543	0.7550	1.0364	1.0263	1.0269
20	0.0265	0.0251	0.0255	0.7494	0.7447	0.7457	1.0217	1.0068	1.0109
				$p = 3.5$					
10	0.0025	0.0019	0.0022	0.8003	0.7939	0.7942	1.1081	1.1066	1.1068
15	0.0081	0.0050	0.0054	0.7796	0.7748	0.7753	1.0587	1.0495	1.0503
20	-0.0133	-0.0121	-0.0122	0.7919	0.7888	0.7888	1.1032	1.0982	1.0982
				$p = 5$					
10	-0.0075	-0.0073	-0.0074	0.8104	0.8086	0.8090	1.0887	1.0830	1.0842
15	0.0030	0.0026	0.0027	0.8007	0.7984	0.7987	1.0761	1.0728	1.0731
20	0.0079	0.0074	0.0074	0.7927	0.7896	0.7901	1.0419	1.0354	1.0363
				$p = 10$					
10	-0.0361	-0.0354	-0.0358	0.8344	0.8337	0.8338	1.1157	1.1136	1.1141
15	-0.0137	-0.0133	-0.0134	0.8106	0.8098	0.8099	1.0588	1.0575	1.0578
20	-0.0051	-0.0047	-0.0048	0.8009	0.8006	0.8007	1.0393	1.0384	1.0385

Simulation results given in Table 2 show that the *ML* estimator of the missing value *m* has the best performance among the others with respect to the *bias*, *MAD* and *MSD* criteria for all cases. It is followed by the *MML* estimator. It should be realized that performances of the *ML* and the *MML* estimators are very close to each other especially for $p \geq 5$. *LS* estimator of *m* has the lowest efficiency for all the sample sizes and the shape parameters as expected. Since, it is well known that the *LS* estimators are optimal if and if only the error terms have normal distribution. However, they lose their efficiency when the distribution of the error terms is non-normal or there exists outliers in the data. It should be noted that *bias*, *MAD* and *MSD* values are very close to each other for all estimators (i.e, *LS*, *ML* and *MML*) when $p = 10$, since $LTS(p = 10)$ is very similar to the normal distribution in shape.

5.2. Effect of Imputation on the Performances of the Estimators of the Model Parameters

We compare the performances of the *LS*, the *ML* and the *MML* estimators of the model parameters (μ_i and σ) given below

$$\hat{\mu}_{i,LS} = \hat{y}_{i.},\ \mu_{i,ML}^{(t)} = \frac{1}{\omega^{(t-1)}}\sum_{j=1}^{n} w_{ij}^{(t-1)} y_{ij},\ \hat{\mu}_{i,MML} = \frac{1}{w}\sum_{j=1}^{n}\beta_j y_{i(j)}$$

and

$$\hat{\sigma}_{LS} = \sqrt{\frac{\sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij}-\hat{\mu}_{i,LS}\right)^2}{N-a}},\ \sigma_{ML}^{(t)} = \sqrt{\frac{\sum_{i=1}^{a}\sum_{j=1}^{n} w_{ij}^{(t-1)}\left(y_{ij}-\mu_{i,ML}^{(t)}\right)^2}{N-a}},\ \hat{\sigma}_{MML} = \frac{B+\sqrt{B^2+4NC}}{2\sqrt{N(N-a)}}$$

where $w = \sum_{j=1}^{n}\beta_j$, $B_i = \frac{2p}{q}\sum_{j=1}^{n}\alpha_j\, y_{i(j)}$, $C_i = \frac{2p}{q}\sum_{j=1}^{n}\beta_j\left(y_{i(j)} - \hat{\mu}_{i,MML}\right)^2$, $B = \sum_{i=1}^{a} B_i$, $C = \sum_{i=1}^{a} C_i$ and $N = an$, see Tiku and Akkaya [43]. Definitions of α_j, β_j and $t_{(j)}$ are given in section 4.

In this section, our aim is to evaluate the effect of "imputation" on the efficiencies of the estimators of the model parameters. For this purpose, we obtain the simulated *mean* and *MSD* values for the *LS*, the *ML* and the *MML* estimators of the model parameters. Simulation results are given in Table 3.

In Table 3, "full" line represents that there is no missing value. Similarly, "$\hat{m}_{LS}$", "$\hat{m}_{ML}$" and "$\hat{m}_{MML}$" lines show that estimators of the model parameters (μ_i, σ) are computed by imputing the estimate of missing value instead of the true value of it. Last column of Table 3 shows the Total *MSD* (*TMSD*) values. *TMSD* criterion is used to compare the joint efficiencies of $\hat{\mu}_i$ and $\hat{\sigma}$. It is defined as shown below

$$TMSD(\hat{\mu}_i, \hat{\sigma}) = MSD(\hat{\mu}_i,) + MSD(\ \hat{\sigma}).$$

It should be noted that small values of $TMSD(\hat{\mu}_i,\ \hat{\sigma})$ implies the joint efficiency of $(\hat{\mu}_i, \hat{\sigma})$. According to the results given in Table 3, we can make the following comments about the efficiencies of $\hat{\mu}_i$ and $\hat{\sigma}$, and also the joint efficiency of $(\hat{\mu}_i, \hat{\sigma})$.

i. *Comparisons for* $\hat{\mu}_i$: It should be realized from Table 3 that efficiencies of $\hat{\mu}_{i,LS}$, $\hat{\mu}_{i,ML}$ and $\hat{\mu}_{i,MML}$ are maximized when we impute the $\hat{m}_{ML}$ instead of the missing value for all cases. Since, its *MSD* values are the closest to the *MSD* values corresponding to the full data. However, when $p \geq 5$, using $\hat{m}_{LS}$, $\hat{m}_{ML}$ or $\hat{m}_{MML}$ do not make any difference in terms of the efficiencies of the estimators of μ_i. It should also be realized that *MSD* values obtained from *ML* and *MML* based imputations are very close to each other when *p* is large. However, their values are much different than the *MSD* values obtained from the *LS* based imputation when *p* is small.

ii. *Comparisons for* $\hat{\sigma}$: Efficiencies of $\hat{\sigma}_{LS}$, $\hat{\sigma}_{ML}$ and $\hat{\sigma}_{MML}$ are more or less the same for each imputed value of *m* except that $p = 2$, 2.5 and $n = 10$. In these cases, $\hat{\sigma}_{LS}$ based on $\hat{m}_{LS}$ gives *MSD* values close to the *MSD* values given in the "full" line. Similar statements can also be made for the $\hat{\sigma}_{ML}$ based on $\hat{m}_{ML}$ and $\hat{\sigma}_{MML}$ based on $\hat{m}_{MML}$.

iii. *Comparisons about the joint efficiencies of* $(\hat{\mu}_i,\ \hat{\sigma})$: Simulation results show that $\hat{\mu}_i$ and $\hat{\sigma}$ are jointly most efficient estimators when we impute $\hat{m}_{ML}$ in the computation of them, especially for small values of *p*. Using $\hat{m}_{ML}$ or $\hat{m}_{MML}$ does not make any difference on the joint efficiency of $\hat{\mu}_i$ and $\hat{\sigma}$ when *p* gets larger. $\hat{m}_{LS}$ can also be used when $p = 10$, since *TMSD* values of $\hat{\mu}_i$ and $\hat{\sigma}$ are very close to each other regardless of looking which imputed value ($\hat{m}_{LS}$, $\hat{m}_{ML}$ or $\hat{m}_{MML}$) instead of *m*.

Table 3. Simulated values of Mean, *MSD* and *TMSD* of the *LS*, *ML* and *MML* estimators of the parameters μ_i and σ

$p=2$	$\hat{\mu}_i$		n=10				n=15				n=20		
		full	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	full	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	full	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$
Mean	LS	-0.0043	-0.0057	-0.0056	-0.0057	-0.0037	-0.0041	-0.0037	-0.0038	0.0012	0.0031	0.0029	0.0030
	ML	-0.0029	-0.0047	-0.0047	-0.0048	0.0019	0.0005	0.0011	0.0009	-0.0004	0.0004	0.0002	0.0003
	MML	-0.0033	-0.0051	-0.0050	-0.0051	0.0003	-0.0005	-0.0001	-0.0002	-0.0003	0.0009	0.0007	0.0007
MSD	LS	0.0954	0.1064	0.0975	0.1004	0.0657	0.0711	0.0667	0.0679	0.0504	0.0533	0.0509	0.0515
	ML	0.0530	0.0606	0.0595	0.0597	0.0352	0.0380	0.0377	0.0377	0.0262	0.0281	0.0278	0.0278
	MML	0.0577	0.0675	0.0631	0.0647	0.0375	0.0415	0.0397	0.0403	0.0282	0.0307	0.0296	0.0299
Mean	LS	0.9335	0.9332	0.9337	0.9335	0.9460	0.9453	0.9456	0.9455	0.9470	0.9466	0.9467	0.9466
	ML	1.0496	1.0549	1.0508	1.0518	1.0318	1.0345	1.0324	1.0328	1.0207	1.0221	1.0209	1.0211
	MML	1.1764	1.1707	1.1700	1.1698	1.1246	1.1202	1.1198	1.1197	1.0929	1.0892	1.0889	1.0888
MSD	LS	0.1298	0.1331	0.1334	0.1333	0.0911	0.0923	0.0924	0.0924	0.0806	0.0814	0.0814	0.0814
	ML	0.0323	0.0341	0.0334	0.0335	0.0203	0.0209	0.0207	0.0207	0.0144	0.0146	0.0145	0.0146
	MML	0.1236	0.1246	0.1239	0.1238	0.0595	0.0593	0.0589	0.0589	0.0373	0.0371	0.0369	0.0369
	$(\hat{\mu}_i,\hat{\sigma})$												
TMSD	LS	0.2252	0.2395	0.2309	0.2337	0.1568	0.1634	0.1591	0.1603	0.1310	0.1347	0.1323	0.1329
	ML	0.0853	0.0947	0.0929	0.0932	0.0555	0.0589	0.0584	0.0584	0.0406	0.0427	0.0423	0.0424
	MML	0.1813	0.1921	0.1870	0.1885	0.0970	0.1008	0.0986	0.0992	0.0655	0.0678	0.0665	0.0668
p=2.5	$\hat{\mu}_i$		n=10				n=15				n=20		
Mean	LS	0.0022	0.0027	0.0026	0.0027	0.0004	0.0019	0.0018	0.0018	0.0016	0.0011	0.0012	0.0012
	ML	0.0019	0.0021	0.0019	0.0021	0.0002	0.0005	0.0003	0.0004	0.0023	0.0021	0.0022	0.0022
	MML	0.0023	0.0026	0.0024	0.0025	0.0001	0.0011	0.0009	0.0010	0.0021	0.0018	0.0019	0.0018
MSD	LS	0.0980	0.1085	0.1033	0.1054	0.0653	0.0706	0.0681	0.0690	0.0499	0.0526	0.0512	0.0516
	ML	0.0734	0.0821	0.0815	0.0816	0.0477	0.0516	0.0513	0.0514	0.0362	0.0385	0.0384	0.0384
	MML	0.0775	0.0878	0.0848	0.0861	0.0500	0.0547	0.0535	0.0539	0.0378	0.0405	0.0398	0.0401
	$\hat{\sigma}$												
Mean	LS	0.9701	0.9693	0.9696	0.9694	0.9814	0.9818	0.9820	0.9819	0.9851	0.9849	0.9850	0.9849
	ML	1.0306	1.0335	1.0316	1.0321	1.0219	1.0235	1.0227	1.0229	1.0182	1.0190	1.0185	1.0186
	MML	1.1267	1.1221	1.1220	1.1218	1.0962	1.0934	1.0933	1.0933	1.0779	1.0753	1.0752	1.0752
MSD	LS	0.0489	0.0498	0.0499	0.0499	0.0421	0.0427	0.0428	0.0427	0.0294	0.0297	0.0297	0.0297
	ML	0.0256	0.0268	0.0266	0.0267	0.0159	0.0163	0.0162	0.0162	0.0119	0.0122	0.0122	0.0122
	MML	0.0603	0.0603	0.0602	0.0601	0.0376	0.0375	0.0374	0.0374	0.0237	0.0236	0.0235	0.0235
	$(\hat{\mu}_i,\hat{\sigma})$												
TMSD	LS	0.1469	0.1583	0.1532	0.1553	0.1074	0.1133	0.1109	0.1117	0.0793	0.0823	0.0809	0.0813
	ML	0.0990	0.1089	0.1081	0.1083	0.0636	0.0679	0.0675	0.0676	0.0481	0.0507	0.0506	0.0506
	MML	0.1378	0.1481	0.1450	0.1462	0.0876	0.0922	0.0909	0.0913	0.0615	0.0641	0.0633	0.0636
p=3	$\hat{\mu}_i$		n=10				n=15				n=20		
Mean	LS	0.0021	0.0011	0.0011	0.0010	0.0039	0.0047	0.0046	0.0046	0.0003	0.0017	0.0016	0.0016
	ML	0.0019	0.0011	0.0011	0.0011	0.0025	0.0027	0.0025	0.0025	-0.0004	0.0003	0.0002	0.0002
	MML	0.0015	0.0008	0.0008	0.0008	0.0027	0.0032	0.0030	0.0031	-0.0002	0.0007	0.0006	0.0006
MSD	LS	0.0975	0.1089	0.1054	0.1060	0.0664	0.0709	0.0692	0.0699	0.0516	0.0549	0.0539	0.0543
	ML	0.0814	0.0910	0.0905	0.0905	0.0540	0.0579	0.0577	0.0578	0.0424	0.0450	0.0449	0.0449
	MML	0.0823	0.0922	0.0911	0.0914	0.0563	0.0606	0.0597	0.0601	0.0438	0.0468	0.0463	0.0465
	$\hat{\sigma}$												
Mean	LS	0.9825	0.9821	0.9823	0.9823	0.9878	0.9879	0.9880	0.9879	0.9905	0.9906	0.9907	0.9906
	ML	1.0262	1.0279	1.0268	1.0269	1.0164	1.0174	1.0169	1.0170	1.0131	1.0138	1.0135	1.0136
	MML	1.0667	1.0613	1.0609	1.0609	1.0768	1.0744	1.0744	1.0743	1.0620	1.0601	1.0601	1.0601
MSD	LS	0.0352	0.0361	0.0361	0.0361	0.0243	0.0248	0.0248	0.0248	0.0184	0.0186	0.0186	0.0186
	ML	0.0240	0.0249	0.0248	0.0248	0.0149	0.0152	0.0152	0.0152	0.0113	0.0114	0.0114	0.0114
	MML	0.0323	0.0323	0.0322	0.0321	0.0267	0.0267	0.0267	0.0266	0.0185	0.0184	0.0184	0.0184
	$(\hat{\mu}_i,\hat{\sigma})$												
TMSD	LS	0.1327	0.1450	0.1415	0.1421	0.0907	0.0957	0.0940	0.0947	0.0700	0.0735	0.0725	0.0729
	ML	0.1054	0.1159	0.1153	0.1153	0.0689	0.0731	0.0729	0.0730	0.0537	0.0564	0.0563	0.0563
	MML	0.1146	0.1245	0.1233	0.1235	0.0830	0.0873	0.0864	0.0867	0.0623	0.0652	0.0647	0.0649
p=3.5	$\hat{\mu}_i$		n=10				n=15				n=20		
Mean	LS	0.0002	0.0005	0.0004	0.0005	0.0048	0.0053	0.0051	0.0051	-0.0060	-0.0067	-0.0066	-0.0066
	ML	-0.0001	0.0000	-0.0004	-0.0004	0.0021	0.0025	0.0022	0.0022	-0.0044	-0.0056	-0.0055	-0.0055
	MML	0.0004	0.0002	0.0002	0.0002	0.0023	0.0028	0.0025	0.0025	-0.0045	-0.0057	-0.0056	-0.0056
MSD	LS	0.0970	0.1084	0.1061	0.1066	0.0663	0.0717	0.0706	0.0707	0.0493	0.0516	0.0509	0.0510
	ML	0.0862	0.0967	0.0964	0.0964	0.0581	0.0629	0.0628	0.0628	0.0424	0.0444	0.0444	0.0444
	MML	0.0872	0.0978	0.0971	0.0974	0.0584	0.0683	0.0630	0.0631	0.0426	0.0447	0.0446	0.0447
	$\hat{\sigma}$												
Mean	LS	0.9868	0.9865	0.9867	0.9866	0.9908	0.9907	0.9908	0.9908	0.9914	0.9910	0.9910	0.9910
	ML	1.0238	1.0251	1.0244	1.0245	1.0152	1.0158	1.0155	1.0155	1.0107	1.0106	1.0104	1.0105
	MML	1.0597	1.0555	1.0553	1.0552	1.0423	1.0392	1.0390	1.0390	1.0316	1.0286	1.0286	1.0286
MSD	LS	0.0267	0.0273	0.0273	0.0273	0.0192	0.0196	0.0197	0.0197	0.0129	0.0130	0.0130	0.0130
	ML	0.0221	0.0228	0.0227	0.0227	0.0141	0.0144	0.0144	0.0144	0.0101	0.0102	0.0102	0.0102
	MML	0.0248	0.0283	0.0283	0.0283	0.0173	0.0173	0.0173	0.0173	0.0115	0.0114	0.0114	0.0114
	$(\hat{\mu}_i,\hat{\sigma})$												
TMSD	LS	0.1237	0.1357	0.1334	0.1339	0.0855	0.0913	0.0903	0.0904	0.0622	0.0646	0.0639	0.0640
	ML	0.1083	0.1195	0.1191	0.1191	0.0722	0.0773	0.0772	0.0772	0.0525	0.0546	0.0546	0.0546
	MML	0.1120	0.1261	0.1254	0.1257	0.0757	0.0856	0.0803	0.0804	0.0541	0.0561	0.0560	0.0561

Table 3 (continued)

p=5	$\hat{\mu}_i$	*n*=10				*n*=15				*n*=20			
		full	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	full	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	full	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$
Mean	LS	0.0091	0.0084	0.0084	0.0084	-0.0008	-0.0006	-0.0006	-0.0006	-0.0037	-0.0033	-0.0033	-0.0033
	ML	0.0093	0.0085	0.0085	0.0085	-0.0012	-0.0010	-0.0011	-0.0010	-0.0040	-0.0037	-0.0037	-0.0037
	MML	0.0094	0.0085	0.0085	0.0085	-0.0011	-0.0009	-0.0009	-0.0009	-0.0041	-0.0037	-0.0037	-0.0037
MSD	LS	0.0997	0.1120	0.1109	0.1112	0.0665	0.0716	0.0711	0.0712	0.0502	0.0531	0.0528	0.0529
	ML	0.0942	0.1064	0.1063	0.1063	0.0627	0.0677	0.0677	0.0677	0.0470	0.0496	0.0495	0.0495
	MML	0.0949	0.1073	0.1069	0.1071	0.0631	0.0682	0.0681	0.0682	0.0473	0.0499	0.0498	0.0499
	$\hat{\sigma}$												
Mean	LS	0.9879	0.9878	0.9879	0.9879	0.9934	0.9933	0.9933	0.9933	0.9951	0.9951	0.9951	0.9951
	ML	1.0117	1.0125	1.0122	1.0123	1.0088	1.0091	1.0091	1.0090	1.0070	1.0072	1.0072	1.0072
	MML	1.0367	1.0341	1.0341	1.0341	1.0292	1.0271	1.0271	1.0271	1.0238	1.0222	1.0222	1.0222
MSD	LS	0.0205	0.0212	0.0212	0.0212	0.0136	0.0138	0.0138	0.0138	0.0097	0.0099	0.0099	0.0099
	ML	0.0191	0.0199	0.0198	0.0198	0.0121	0.0123	0.0123	0.0123	0.0087	0.0089	0.0089	0.0089
	MML	0.0221	0.0226	0.0226	0.0226	0.0139	0.0140	0.0140	0.0140	0.0097	0.0098	0.0098	0.0098
	$(\hat{\mu}_i, \hat{\sigma})$												
TMSD	LS	0.1202	0.1332	0.1321	0.1324	0.0801	0.0854	0.0849	0.0850	0.0599	0.0630	0.0627	0.0628
	ML	0.1133	0.1263	0.1261	0.1261	0.0748	0.0800	0.0800	0.0800	0.0557	0.0585	0.0584	0.0584
	MML	0.1170	0.1299	0.1295	0.1297	0.0770	0.0822	0.0821	0.0822	0.0570	0.0597	0.0596	0.0597
p=10	$\hat{\mu}_i$	*n*=10				*n*=15				*n*=20			
Mean	LS	0.0014	-0.0023	-0.0022	-0.0022	0.0010	0.0001	0.0001	0.0001	-0.0006	-0.0008	-0.0008	-0.0008
	ML	0.0021	-0.0016	-0.0015	-0.0016	0.0014	0.0004	0.0005	0.0005	-0.0003	-0.0005	-0.0005	-0.0005
	MML	0.0017	-0.0019	-0.0018	-0.0019	0.0014	0.0004	0.0004	0.0004	-0.0003	-0.0005	-0.0005	-0.0005
MSD	LS	0.0997	0.1109	0.1107	0.1107	0.0666	0.0718	0.0717	0.0718	0.0506	0.0536	0.0535	0.0535
	ML	0.0989	0.1098	0.1098	0.1098	0.0659	0.0711	0.0711	0.0711	0.0501	0.0529	0.0529	0.0529
	MML	0.0989	0.1099	0.1098	0.1099	0.0660	0.0711	0.0711	0.0711	0.0501	0.0530	0.0530	0.0530
	$\hat{\sigma}$												
Mean	LS	0.9912	0.9909	0.9909	0.9909	0.9953	0.9953	0.9953	0.9953	0.9967	0.9967	0.9967	0.9967
	ML	1.0021	1.0022	1.0021	1.0021	1.0028	1.0029	1.0029	1.0029	1.0022	1.0023	1.0023	1.0023
	MML	1.0138	1.0125	1.0125	1.0125	1.0126	1.0118	1.0118	1.0118	1.0107	1.0101	1.0100	1.0100
MSD	LS	0.0162	0.0167	0.0167	0.0167	0.0103	0.0106	0.0106	0.0106	0.0079	0.0080	0.0080	0.0080
	ML	0.0161	0.0166	0.0166	0.0166	0.0102	0.0104	0.0104	0.0104	0.0077	0.0079	0.0079	0.0079
	MML	0.0169	0.0172	0.0172	0.0172	0.0106	0.0108	0.0108	0.0108	0.0081	0.0082	0.0082	0.0082
	$(\hat{\mu}_i, \hat{\sigma})$												
TMSD	LS	0.1159	0.1276	0.1274	0.1274	0.0769	0.0824	0.0823	0.0824	0.0585	0.0616	0.0615	0.0615
	ML	0.1150	0.1264	0.1264	0.1264	0.0761	0.0815	0.0815	0.0815	0.0578	0.0608	0.0608	0.0608
	MML	0.1158	0.1271	0.1270	0.1271	0.0766	0.0819	0.0819	0.0819	0.0582	0.0612	0.0612	0.0612

6. AN APPLICATION: PEAK DISCHARGE DATA

In this section, we use the data set to illustrate and compare the performances of the estimators of the missing value. The data set was taken from Montgomery [26]. In the data set, four different methods of estimating flood flow frequency are used to determine whether they produce equivalent estimates of peak discharge in case of application to the same watershed. Each procedure is used six times on the watershed, and the resulting discharge data (in cubic feet per second) are given in the Table 4.

Table 4. The peak discharge data

Treatments			
I	II	III	IV
0.34	0.91	6.31	17.15
0.12	2.94	8.37	11.82
1.23	2.14	9.75	10.95
0.70	2.36	6.09	17.20
1.75	2.86	9.82	14.35
0.12	4.55	7.24	16.82

To model peak discharge data, we use one-way *ANOVA* given below

$$y_{ij} = \mu_i + \varepsilon_{ij},\ i = 1,2,3,4;\ j = 1,2,\cdots,6$$

Before analyzing the data set, we identify the shape parameter p by using the Quantile-Quantile (Q-Q) plot which is the graphical technique to check the validity of a distributional assumption for the peak discharge data. We plot the order statistics of

$$\hat{\varepsilon}_{ij} = y_{ij} - \hat{\mu}_i,\ i = 1,2,3,4;\ j = 1,2,\cdots,6$$

against the quantiles of the normal distribution, see Figure 2. We obtain the Q-Q plot of the peak discharge data for various different values of the shape parameter p of *LTS* distribution. It is seen that $LTS(p = 2.3)$ provides better fitting to the peak discharge data than the normal distribution, since the data points in the $LTS(p = 2.3)$ Q-Q plot do not deviate too much from a straight line, see Figure 3.

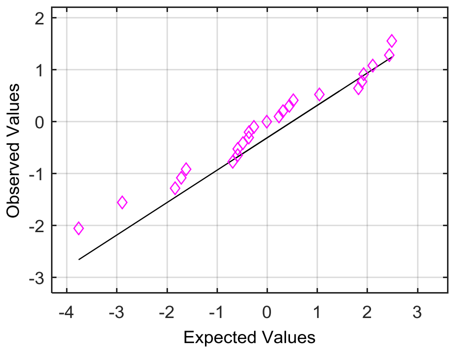

Figure 2. Normal distribution Q-Q plot of the peak discharge data

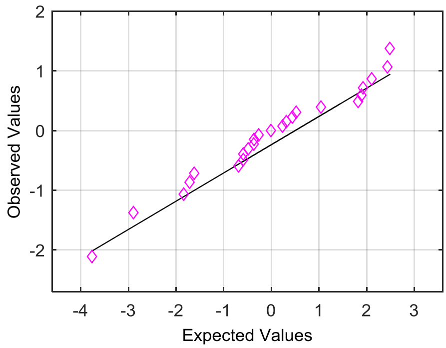

Figure 3. $LTS(p = 2.3)$ distribution Q-Q plot of the peak discharge data

The Akaike Information Criterion (*AIC*) values based on *ML* estimates are also calculated for the *LTS* (when $p = 2.3$) and the normal distribution. The *AIC* values are obtained as follows

$$AIC_{LTS} = 52.3080 \text{ and } AIC_N = 53.8291,$$

respectively. It is clear that *LTS* distribution provides better fitting to the data than the normal distribution according to the *AIC* value.

In this example, there is no missing value in the data, therefore we create a missing value in the second treatment (i.e, $m = y_{23}$) according to the *MCAR* mechanism by using the methodology proposed by Brand et al. [7]. Then, we obtain the estimates of *m* based on *LS*, *ML* and *MML* methodologies. *Bias* values corresponding to them, are also given in Table 5.

The results given in Table 5 show that the *ML* estimator of missing value has the smallest *bias* among the others. This result is consistent with the simulation results given in Table 2.

Table 5. The *LS*, the *ML* and the *MML* estimates of the missing value for the data

$m = y_{23}$	$\hat{m}_{LS}$	$\hat{m}_{ML}$	$\hat{m}_{MML}$	$Bias_{LS}$	$Bias_{ML}$	$Bias_{MML}$
2.1400	2.5640	2.5050	2.5110	0.1798	0.1333	0.1377

7. CONCLUSIONS

In this paper, we consider the problem of estimating the missing value in one-way *ANOVA* when the error terms have *LTS* distribution. Estimates of missing value are obtained by using the *LS*, the *ML* and the *MML* methodologies. The *EM* algorithm is used for the *ML* estimate of the missing value. The performances of the estimates are compared according to the three different criteria. The results of simulation study show that the *ML* and the *MML* estimators of the missing value are more efficient than the corresponding *LS* estimator with respect to the *bias*, *MAD* and the *MSD* criteria. It should also be noted that the *LS* estimator did not perform well in accordance with all comparison measures for all the sample sizes and the shape parameters, because it is very sensitive to the deviations from the normal distribution. We also examined how estimates of the missing value affect the performances of the estimators of the model parameters. Estimates of missing value are used to fill in the missing value, so that incomplete data becomes complete and then we obtain the estimates of the model parameters μ_i and σ by using the *LS*, the *ML* and the

MML methodologies. Finally, we compare the efficiencies of the estimators of the model parameters based on "full data" and "imputed missing data". The *ML* estimate of missing value in one-way *ANOVA* is the best among the others since the estimates of the model parameters based on *ML* estimate of the missing value are numerically very close to the estimates of the model parameters based on full data. It should be noted that the *MML* estimator is non-iterative and numerically very close to the *ML* estimator for estimating the missing value. The results given in this paper can easily be extended to more than one missing value case and to the hypothesis testing problems.

REFERENCES

[1] Allan F.G. and Wishart J., (1930) A Method of Estimating the Yield of a Missing Plot in Field Experiments, *Journal of Agricultural Science* 20, 399-406.

[2] Anderson R.L., (1946) Missing-plot Techniques, Biometrics Bulletin 2, 41–47.

[3] Aydin D., (2013) Robust Estimation of Missing Observations in Experimental Design, PhD Thesis, *Graduate School of Natural and Applied* Sciences, *Ankara University*, Ankara Turkey (In Turkish).

[4] Aydin D. and Senoglu B., (2017) A Comparison of Different Methods for Estimating the Missing Value in Two-Way ANOVA, *International Journal of Mathematics and Statistics* 18, 47–63.

[5] Bartlett M.S., (1937) Some Examples of Statistical Methods of Research in Agriculture And Applied Biology, *Journal of the Royal Statistical Society-Supplement* 4, 137–170,

[6] Bickel P.J. and Doksum K.A., (1981) An Analysis of Transformation Revisited, *Journal of American Statistical Association* 76, 296–311.

[7] Brand J.P.L. and Van Buuren S., Groothuis-Oudshoorn K. and Gelsema E.S., (2003) A toolkit in SAS for the evaluation of multiple imputation methods, *Statistica Neerlandica* 57, 36–45.

[8] Box G.E.P. and Cox D.R., (1964) An Analysis of Transformations (With Discussion), *Journal of the Royal Statistical Society* 26, 211–252.

[9] Cochran W.G. and Cox G.M., (1957) Experimental Designs, *John Wiley & Sons Inc.*, Canada.

[10] Dempster A.P., Laird N.M. and Rubin D., (1977) Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society* 39, 1–38.

[11] Dodge Y., (1985) Analysis of Experiments with Missing Data, *John Wiley & Sons Inc.*, Canada.

[12] Elveback L.R., Guillier C.L. and Keating F.R., (1970) Health, Normality and the Ghost of Gauss, *Journal of the American Medical Association* 211, 69–75.

[13] Enders C.K., (2010) Applied Missing Data Analysis, *Guilford Press*, USA.

[14] Engels J. M. and Diehr P., (2003) Imputation of Missing Longitudinal Data: a Comparison of Methods, *Journal of Clinical Epidemiology* 56, 968–976.

[15] Geary R.C., (1947) Testing for Normality, *Biometrika* 34, 209–242.

[16] Hartley H.O., (1956) Programming Analysis of Variance for General Purpose *Computers, Biometrics* 12, 110–122.

[17] Healy M.J.R. and Westmacott M., (1956) Missing Values in Experiments Analyzed On Automatic Computers, *Journal of Applied Statistics* 5, 203–206.

[18] Huber P.J., (1964) Robust Estimation of a Location Parameter, *The Annals of Mathematical Statistics* 35, 73–101.

[19] Huber P.J., (1981) Robust Statistics, *John Wiley & Sons Inc*., New York, USA.

[20] Jarrett R.G., (1978) The Analysis of Designed Experiments with Missing Observations, *Journal of Applied Statistics* 27, 38–46.

[21] Islam M.Q. and Tiku M.L., (2005), Multiple Linear Regression Model Under Non-normality, *Communications in Statistics-Theory and Methods* 33, 2443–2467.
[22] Lange, K., Little, R. and Taylor, J., (1989) Robust Statistical Modeling Using the t Distribution, *Journal of the American Statistical Association* 84, 881–896.
[23] Little, R.J.A. and Rubin, D.B., (1983) On Jointly Estimating Parameters and Missing Data by Maximizing the Complete-Data Likelihood, *American Statistician* 37, 218–220.
[24] Liu C. and Rubin D.B., (1995) ML Estimation of the t Distribution Using EM and its Extensions, ECM and ECME, *Statistica Sinica* 5, 19–39.
[25] Mclachlan G.J. and Krishnan T., (2007) The EM Algorithm and Extensions, *John Wiley & Sons Inc*., New Jersey, USA.
[26] Montgomery D.C., (2005) Design and Analysis of Experiments, *John Wiley & Sons Inc*., USA.
[27] Mutan O.C. and Şenoğlu B., (2009) A Monte Carlo Comparison of Regression Estimators when the Error Distribution is Long-Tailed Symmetric, *Journal of Modern Applied Statistical Methods* 8(1), 126–137.
[28] Pearson E.S., (1931) The Analysis of Variance in Cases of Non-normal Variation, *Biometrika* 23, 114–13.
[29] Peel D. and Mclachlan G.J., (2000) Robust Mixture Modelling Using the t Distribution, *Statistics and Computing* 10, 339–348.
[30] Preece D.A., (1971) Iterative Procedures for Missing Values in Experiments, *Technometrics* 13, 743–753.
[31] Rao C.R. and Toutenburg H., (1999) Linear Models: Least Square and Alternatives, *Springer*, New York.
[32] Reynolds J.F., (1995) Estimating Missing Observations in ANOVA, *International Journal of Mathematical Education in Science and Technology* 26, 895–901.
[33] Rubin D.B., (1972) A Non-Iterative Algorithm for Least Squares Estimation of Missing Values in any Analysis of Variance Design, *Journal of the Royal Statistical Society* 21, 136–141.
[34] Rubin D.B., (1976) Inference and Missing Data, *Biometrika* 63, 581–592.
[35] Senoglu, B. and Avcioglu, M.D., (2009) Analysis of Covariance with Non-normal Errors, *International Statistical Review* 77, 3, 366–377.
[36] Şenoğlu B. and Acıtaş Ş., (2016) Robust Factorial ANCOVA with LTS Error Distributions, *Hacettepe Journal of Mathematics and Statistics* Doi: 10.15672/HJMS.201612918797.
[37] Subramani J. and Ponnuswamy K.N., (1989) A Non-Iterative Least Squares Estimation of Missing Values in Experimental Designs, *Journal of Applied Statistics* 16, 77–86.
[38] Subramani J., (1993) Non-iterative Least Squares Estimation of Missing Values in Hyper-Graeco-Latin Square Designs, *Biometrical Journal* 35, 465–470.
[39] Srivastava A.B.L., (1959) Effect of Non-Normality on The Power of the Analysis of Variance Test, *Biometrika* 46, 114–122.
[40] Tiku M.L., (1967) Estimating the Mean and Standard Deviation from a Censored Normal Sample, *Biometrika* 54, 155–165.
[41] Tiku M.L. and Kumra S., (1981) Expected values and variances and covariances of order statistics for a family of symmetric distributions (Student's t), Selected Tables in Mathematical Statistics 8, 141–270 American Mathematical Society, Providence RI, USA.
[42] Tiku M.L., Islam M.Q. and Selcuk S.A., (2001) Non-normal Regression, Part II: Symmetric Distributions, *Communications in Statistics - Theory and Methods* 30, 1021–1045.
[43] Tiku M.L. and Akkaya A.D., (2004) Robust Estimation and Hypothesis Testing, *New Age International (P) Limited Publishers*, New Delhi, India.

[44] Wilkinson G.N., (1958) Estimation of Missing Values for the Analysis of Incomplete Data, *Biometrics* 14, 174–194.
[45] Yates F., (1933) The Analysis of Replicated Experiments When the Field Results are Incomplete, *Empire Journal of Experimental Agriculture* 1, 129–142.