Revista Colombiana de Estadística
January 2015, Volume 38, Issue 1, pp. 75 to 91
DOI: http://dx.doi.org/10.15446/rce.v38n1.48802

Estimation and Testing in One-Way ANOVA when the Errors are Skew-Normal

Estimación y pruebas de hipótesis en ANOVA a una vía cuando los errores se distribuyen como normal sesgados

Nuri Celik[1,a], Birdal Senoglu[2,b], Olcay Arslan[2,c]

[1]Department of Statistics, Science Faculty, Bartin University, Bartin, Turkey
[2]Department of Statistics, Science Faculty, Ankara University, Ankara, Turkey

Abstract

We consider one-way analysis of variance (ANOVA) model when the error terms have skew- normal distribution. We obtain the estimators of the model parameters by using the maximum likelihood (ML) and the modified maximum likelihood (MML) methodologies (see, Tiku 1967). In the ML method, iteratively reweighting algorithm (IRA) is used to solve the likelihood equations. The MML approach is a non-iterative method used to obtain the explicit estimators of model parameters. We also propose new test statistics based on these estimators for testing the equality of treatment effects. Simulation results show that the proposed estimators and the tests based on them are more efficient and robust than the corresponding normal theory solutions. Also, real data is analysed to show the performance of the proposed estimators and the tests.

***Key words*:** ANOVA, Modified Likelihood, Iteratively Reweighting Algorithm, Skew-Normal, Monte Carlo Simulation, Robustness.

Resumen

Se considera el modelo de análisis de varianza a una vía (ANOVA) cuando los términos de error siguen una distribución normal sesgada. Se obtienen estimadores de los parámetros desconocidos mediante el uso de la metodología de máxima verosimilitud (ML). Se proponen nuevos estadísticos de prueba basados en estos estimadores. Los resultados de la simulación muestran que los estimadores propuestos y los tests basados en ellos son más eficientes y robustos que los correspondientes a las soluciones de la teoría normal. Un conjunto de datos real es analizado con el fin de mostrar el desempeño de los estimadores propuestos y sus tests relacionados.

***Palabras clave*:** ANOVA, estimación, normal sesgada, pruebas de hipótesis, robustez.

[a]Professor. E-mail: ncelik@bartin.edu.tr
[b]Professor. E-mail: senoglu@science.ankara.edu.tr
[c]Professor. E-mail: oarslan@ankara.edu.tr

1. Introduction

Consider the following one-way ANOVA model,

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, i = 1, 2, \ldots, a; j = 1, 2, \ldots, n \tag{1}$$

where, y_{ij} are the responses corresponding to jth observation in the ith treatment, μ is the overall mean, α_i is the effect of ith treatment and ϵ_{ij} are the independent and identically distributed (iid) random error terms.

In general, normality assumption is made for the random error terms and the well known least squares (LS) method is used for estimating model parameters. However, in the literature, there are numerous studies pointing out that non-normal distributions are more prevalent than normal distribution , in practice, see for example, (Pearson 1932, Geary 1947, Huber 1981, Tan & Tiku 1999). It is known that LS estimators of the parameters and the test statistics based on them lose their efficiency when the normality assumption is not satisfied, (see, Tukey 1960). That is why there is great interest in studying the effect of non-normality on the F statistics used for testing the main effects and the interaction in the framework of experimental design; see, for example, (Geary 1947, Srivastava 1959, Donaldson 1968, Spjotvoll & Aastveit 1980, Tan & Tiku 1999, Senoglu & Tiku 2001). The following conclusions have been drawn from these studies. For numerous non-normal distributions:

i. Type I error of the F statistic is not much different than that for a normal distribution. This is essentially due to the central limit theorem.

ii. Power of the F test is considerably lower than that for a normal distribution. This is essentially due to the inefficiency of the sample mean.

See Senoglu & Tiku (2001) and the references therein. These conclusions are particularly true for non normal distributions having skewness in different directions (Senoglu & Tiku 2002).

Therefore, it is necessary to obtain new F statistics whose distribution provides satisfactory approximations to the percentage points of the null distribution when the distribution of the error terms is non-normal (see condition i). The proposed test should also maintain higher power than the classical F test based on LS estimators (see condition ii).

There are various ways of analyzing non-normal data, such as Box-Cox normalizing transformation and nonparametric methods. However, in this study, we adopt the parametric ML and MML methods where original data are used rather than transformed data. In the ML method, the likelihood equations are solved iteratively by using the iteratively reweighting algorithm (IRA). However, in the MML method, the explicit estimators of model parameters are obtained by approximating the likelihood equations.

In this study, we assume that the distribution of the error terms in one-way ANOVA model in (1) is Azzalini's skew-normal (Azzalini 1985, 1986) and obtain the ML and the MML estimators of the model parameters. We then propose new test statistics based on these estimators. To the best of our knowledge, there is no previous work assuming $SN(\lambda)$ as an error distribution in the context of ANOVA. The reason for choosing the $SN(\lambda)$ as an error distribution is that it includes the normal distribution as well as plausible alternatives thereof with different levels of skewness and kurtosis. Therefore, $SN(\lambda)$ distribution is considered to be an extension of normal distribution. This provides us flexibility for modeling the data with $normal - like$ shape but with skewness and heavy tails. It is also useful for modeling the data having normal distribution with outliers and contamination. Its mathematical tractability is another reason for using $SN(\lambda)$ in this study.

The rest of the paper is organized as follows. In Section 2, $SN(\lambda)$ distribution is introduced. The ML and the MML estimators are derived in Section 3 and Section 4, respectively. Efficiencies of the ML and the MML estimators are compared via Monte Carlo simulation study in Section 5. New test statistics for testing the equality of treatment effects are proposed in Section 6. Power comparisons and robustness properties of these tests are also given in this section. A real life example is analyzed in Section 7 to present the application of the proposed estimators and the tests based on them. Our conclusions are presented.

2. Skew-Normal Distribution

The probability density function (pdf) of the $SN(\lambda)$ distribution is given by

$$h(z) = 2\phi(z)\Phi(\lambda z) \tag{2}$$

where $\phi(z)$ and $\Phi(z)$ are the pdf and the cumulative distribution function (cdf) of the standard normal distribution, respectively. λ is the skewness parameter, it is also known as the shape parameter since it regulates the shape of the distribution. If a random variable Z has a skew-normal distribution with parameter λ then it is denoted by $Z \sim SN(\lambda)$. Some extensions of this distribution con found in Martínez-Flórez, Vergara-Cardozo & González (2013) and Pereira, Marques & da Costa (2012).

It may be noted that for λ=0, $SN(\lambda)$ reduces to the well known standard normal distribution $N(0,1)$. When $\lambda \to \infty$, $SN(\lambda)$ converges to the half-normal distribution, $h(z)$ is strongly unimodal for fixed λ. It is right skewed for $\lambda > 0$ and left skewed for $\lambda < 0$. $SN(\lambda)$ distribution has also the following properties:

i. If $Z \sim SN(\lambda)$ then $-Z \sim SN(-\lambda)$

ii. If $Z \sim SN(\lambda)$ then $Z^2 \sim \chi^2_1$ (see, Azzalini 2005).

To better understand the shape of the $SN(\lambda)$ distribution, see the coefficients of skewness (γ_1) and the kurtosis (γ_2) for some representative values of λ given in Table 1.

It is clear from Table 1 that the skewness of the distribution increases as the skewness parameter λ increases (in absolute value). Skewness of the $SN(\lambda)$ distribution takes values in the interval $(-0.995, 0.995)$ and the maximum value of its kurtosis is 3.869. Here, it should be noted that the skewness values corresponding to the positive λ values are exactly the same, but with opposite sign, with

Table 1: The skewness and the kurtosis values of the $SN(\lambda)$ distribution.

λ	0.0	1.0	2.0	3.0	4.0	5.0	10	20	∞
γ_1	0.00	0.14	0.45	0.67	0.78	0.85	0.96	0.99	0.995
γ_2	3.00	3.06	3.31	3.51	3.63	3.71	3.82	3.86	3.869

the skewness values corresponding to the negative λ values. Therefore, in Table 1, we just reproduce the skewness values corresponding to the positive λ values for brevity. It can also be seen that the $SN(\lambda)$ and the normal distribution are indistinguishable for $\lambda < 3$.

Here and in many other studies, we consider a more general form of the distribution given in (2) by performing a change of location and scale:

$$Y = \mu + \sigma Z \tag{3}$$

Based on this linear transformation, pdf of the random variable Y is obtained as shown below,

$$h(y) = \frac{2}{\sigma}\phi(\frac{y-\mu}{\sigma})\Phi(\lambda\frac{y-\mu}{\sigma}) \tag{4}$$

where, $\mu \in \mathcal{R}$ is the location parameter and $\sigma \in \mathcal{R}^+$ is the scale parameter. If the random variable Y has $SN(\lambda)$ distribution with the parameters μ, σ and λ, then it is denoted by $Y \sim SN(\mu, \sigma, \lambda)$. The expected value and the variance of $SN(\mu, \sigma, \lambda)$ distribution are given by,

$$E(Y) = \mu + \sqrt{\frac{2\lambda^2}{\pi(1+\lambda^2)}}\sigma,\ V(Y) = (1 - \frac{2\lambda^2}{\pi(1+\lambda^2)})\sigma^2 \tag{5}$$

respectively.

3. Maximum Likelihood Estimator

Consider the model (1) and assume the distribution of $\epsilon_{ij}, (i = 1, 2, \dots, a; j = 1, 2, \dots, n)$ is skew-normal $SN(0, \sigma, \lambda)$.

$$h(\epsilon) = \frac{2}{\sigma}\phi(\frac{\epsilon}{\sigma})\Phi(\lambda\frac{\epsilon}{\sigma}), -\infty < \epsilon < \infty \tag{6}$$

Here, it should be noted that the skewness parameter is assumed to be known throughout the study. Since the ML method gives doubtful estimates when we estimate the location, the scale and the shape parameters simultaneously unless large samples ($n > 250$ or so) are available, (see, Bowman & Shenton 2001, Kantar & Senoglu 2008). See also, the Introduction of Acitas, Kasap, Senoglu & Arslan (2013). However, the sample size is much smaller than 250 in the context of experimental design. Therefore, in this study, we only estimate the location and the scale parameters for a better fitting. In spite of the fact that the shape parameter is assumed to be known, in practice, we must identify its value. Shape parameters can be identified by using various techniques, such as Q-Q plots, goodness-of fit tests etc. The algorithm given in Acitas et al. (2013, p. 417) can also be used for the identification of the shape parameter, see also Islam & Tiku (2004). Suppose that the value of the shape parameter in skew-normal distribution might be somewhat misspecified by using these techniques. Then the question arises what effect will it have on the efficiencies of the location and the scale estimators. The answer is that this does not adversely affect the efficiencies of the estimators since the estimators obtained in this study are robust to plausible deviations of the true model.

To obtain the ML estimators of the unknown parameters in model (1), the log-likelihood function

$$\ln L = N\ln 2 - N\ln\sigma - \frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{n} z_{ij}^2 + \frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{n}\ln\Phi(\lambda z_{ij}) \tag{7}$$

is maximized with respect to the unknown parameters μ, α_i and σ. Here $z_{ij} = \frac{\epsilon_{ij}}{\sigma} = \frac{y_{ij}-\mu-\alpha_i}{\sigma}$.

By differentiating the log-likelihood function with respect to the unknown parameters and equating them to zero we obtain the following likelihood equations

$$\begin{aligned}
\frac{\partial \ln L}{\partial \mu} &= \sum_{i=1}^{a}\sum_{j=1}^{n} z_{ij} - \lambda\sum_{i=1}^{a}\sum_{j=1}^{n}\frac{\phi(\lambda z_{ij})}{\Phi(\lambda z_{ij})} = 0\\
\frac{\partial \ln L}{\partial \alpha_i} &= \sum_{j=1}^{n} z_{ij} - \lambda\sum_{j=1}^{n}\frac{\phi(\lambda z_{ij})}{\Phi(\lambda z_{ij})} = 0\\
\frac{\partial \ln L}{\partial \sigma} &= -N + \sum_{i=1}^{a}\sum_{j=1}^{n} z_{ij}^2 - \lambda\sum_{i=1}^{a}\sum_{j=1}^{n} z_{ij}\frac{\phi(\lambda z_{ij})}{\Phi(\lambda z_{ij})} = 0
\end{aligned} \tag{8}$$

Solutions of these equations are the ML estimators. These equations have no explicit solutions; therefore we resort to iterative methods.

If we appropriately reorganize the likelihood equations in (8) and define the weight function w_{ij} as below

$$w_{ij} = \frac{\phi(\lambda z_{ij})}{\Phi(\lambda z_{ij})}$$

the likelihood equations can be written as follows:

$$\hat{\mu} = \bar{y}_{..} - \lambda\bar{w}_{..}\hat{\sigma},\ \hat{\alpha_i} = \bar{y}_{i.} - \bar{y}_{..} - \lambda(\bar{w}_{i.} - \bar{w}_{..})\hat{\sigma},\ \hat{\sigma}^2 = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2}{N(1-\lambda^2 t^2)} \tag{9}$$

where

$$\bar{y}_{i.} = \frac{\sum_{j=1}^{n} y_{ij}}{n}, \bar{y}_{..} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}}{N}, \bar{w}_{i.} = \frac{\sum_{j=1}^{n} w_{ij}}{n}, \bar{w}_{..} = \frac{\sum i = 1^a \sum_{j=1}^{n} w_{ij}}{N}$$

and t=$\frac{\sum_{i=1}^{a} w_{i.}^2}{a}$

We use IRA which is very popular in robustness studies to compute the ML estimates of the parameters. It can be shown that IRA is an expectation-maximization (EM) type algorithm so that its convergence is guaranteed (see, Arslan & Genc 2009). Also Arrellano-Valle, Bolfarine & Lachos (2005),Lachos, Bolfarine, Arellano-Valle & Montenegro (2007), Xie, Wei & Lin (2009), Lachos, Ghosh & Arellano-Valle (2010), Lachos, Bandyopadhyay & Garay (2011), Garay, Lachos & Abanto-Valle (2011), Garay, Lachos, Labra & Ortega (2013). In the above mentioned papers, skew normal is used as an error distribution in the context of regression and linear mixed models. Steps of the IRA are given below.

Iteratively reweighting algorithm (IRA):

i. Identify the initial estimates $\mu_i^{(0)}(i = 1, 2, \ldots, a)$ and $\sigma^{(0)}$ for μ_i and σ, respectively.

ii. Compute the weights $w_{ij}^{(m)} = \frac{\phi(\lambda z_{ij}^{(m)})}{\Phi(\lambda z_{ij}^{(m)})}$, the averages $\bar{w}_{i.}^{(m)}$ and $t^{(m)} = \frac{\sum_{i=1}^{a}(\bar{w}_{i.}^{(m)})^2}{a}$ where $z_{ij}^{(m)} = \frac{y_{ij}-\mu_i^{(m)}}{\sigma^{(m)}}(i = 1, 2, \ldots, a; j = 1, 2, \ldots, n)$. Here, m is the number of iterations and takes the values $1, 2, 3, \ldots$

iii. Find new estimates of the parameters by using the following updating equations $\mu_i^{(m+1)} = \bar{y}_{..} - \lambda\bar{w}_{i.}^{(m)}\sigma^{(m)}$ and $(\sigma^2)^{(m)} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{i_j}-\bar{y}_{i.})^2}{N(1-\lambda^2(t^{(m)})^2)}$

iv. Continue the iterations until $|\mu_i^{(m+1)} - \mu_i^{(m)}| < d$ and $|\sigma^{(m+1)} - \sigma^{(m)}| < d$ where d is a predetermined small constant.

It should be noted that LS estimates are used as initial estimates for this algorithm. However, some other robust estimates can also be used as initial estimates.

4. Modified Maximum Likelihood Estimator

In this section, we use the MML methodology originated by Tiku (1967) to obtain the explicit estimators of the model parameters by approximating the likelihood equations appropriately. This methodology is used to alleviate the computational difficulties encountered in solving the likelihood equations given above. MML methodology proceeds as follows: Let

$$y_{i_{(1)}} < y_{i_{(2)}} < \cdots < y_{i_{(n)}}, i = 1, 2, \ldots, a \tag{10}$$

be the order statistics obtained by arranging $y_{i_j}(i = 1, 2, \ldots, a; j = 1, 2, \ldots, n)$ in ascending order. The likelihood equations in (8) can be written in terms of the order statistics as shown below, since complete sums are invariant to ordering $(i.e \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} y_{(i)})$.

$$\begin{aligned}\frac{\partial \ln L}{\partial \mu} &= \sum_{i=1}^{a}\sum_{j=1}^{n} z_{i_{(j)}} - \lambda \sum_{i=1}^{a}\sum_{j=1}^{n} g(z_{i_{(j)}}) = 0\\ \frac{\partial \ln L}{\partial \alpha_i} &= \sum_{j=1}^{n} z_{i_{(j)}} - \lambda \sum_{j=1}^{n} g(z_{i_{(j)}}) = 0\\ \frac{\partial \ln L}{\partial \sigma} &= -N + \sum_{i=1}^{a}\sum_{j=1}^{n} z_{i_{(j)}}^2 - \lambda \sum_{i=1}^{a}\sum_{j=1}^{n} z_{i_{(j)}} g(z_{i_{(j)}}) = 0\end{aligned} \tag{11}$$

Here, $g(z) = \frac{\phi(\lambda z)}{\Phi(\lambda z)}$ and $z_{i_{(j)}} = \frac{y_{i_{(j)}} - \mu - \alpha_i}{\sigma}$. It should be noted that the last two terms of $\frac{\partial \ln L}{\partial \sigma}$ are obtained by simply multiplying the terms of $\frac{\partial \ln L}{\partial \mu}$ by $z_{i_{(j)}}$. $z_{i_{(j)}}$ is the loading factor and instrumental in yielding an estimator which is always real and positive. Then, we linearize the intractable terms in (11) by using the first two terms of Taylor series expansion around the expected values of the standardized order statistics, i.e., $t_{(j)} = E(z_{i_{(j)}}), j = 1, 2, \dots, n$. This linearization yields

$$g(z_{i(j)}) = \alpha_j - \gamma_j z_{i(j)}, i = 1, 2, \dots, a;\ j = 1, 2 \dots, n \tag{12}$$

where

$$\gamma_j = \frac{\phi(\lambda t_{(j)})}{\Phi(\lambda t_{(j)})} \left(\frac{\lambda^2 t_{(j)} \Phi(\lambda t_{(j)} + \lambda \phi(\lambda t_{(j)})}{\Phi(\lambda t_{(j)})} \right)$$

and

$$\alpha_{(j)} = \frac{\phi(\lambda t_{(j)})}{\Phi(\lambda t_{(j)})} + t_{(j)} \gamma_{(j)}$$

The exact values of $t_{(j)}$ are not available, however, we use their approximate values obtained from the equation,

$$F(t_{(j)}) = \int_{-\infty}^{t_{(j)}} h(z) dz = \frac{j}{n+1}, (j = 1, 2 \dots, n) \tag{13}$$

(see, Tiku & Akkaya 2004). Here, we use the property: If $F(z_j) \sim U(0,1)$ then $F(z_{(j)}) \sim Beta(j, n-j+1)$ with the expected value $\frac{j}{n+1}, (j = 1, 2 \dots, n)$.

Incorporating equation (12) into the likelihood equations in (11), we obtain the modified likelihood equations $\frac{\partial \ln L^*}{\partial \mu}$, $\frac{\partial \ln L^*}{\partial \alpha_i}$ and $\frac{\partial \ln L^*}{\partial \sigma}$. The solutions of these modified likelihood equations are the following MML estimators

$$\hat{\mu} = \hat{\mu}_{..} - \lambda \frac{\Delta}{m} \hat{\sigma},\ \hat{\alpha}_i = \hat{\mu}_{i.} - \hat{\mu}_{..},\ \hat{\sigma} = \frac{B + \sqrt{B^2 - 4NC}}{2\sqrt{N(N-a)}} \tag{14}$$

where

$$\hat{\mu}_{i.} = \frac{\sum_{j=1}^{n} \beta_{(j)} y_{i(j)}}{m}, \hat{\mu}_{..} = \frac{\sum_{i=1}^{a} \hat{\mu}_{i.}}{a}, \Delta = \lambda \sum_{j=1}^{n} \alpha_{(j)}, \beta_{(j)} = 1 + \lambda \gamma_{(j)}, m = \sum_{j=1}^{n} \beta_{(j)}$$

$$B=\sum_{i=1}^{a}\sum_{j=1}^{n}\alpha_{(j)}(y_{i(j)}-\hat{\mu}_{i.}),\ C=\sum_{i=1}^{a}\sum_{j=1}^{n}\beta_{(j)}(y_{i(j)}-\hat{\mu}_{i.})^2$$

The divisor N in the expression for $\hat{\sigma}$ was replaced by $\sqrt{N(N-a)}$ as a bias correction. MML estimators have the following properties:

i. They are the functions of sample observations and are easy to compute.

ii. They are asymptotically equivalent to the ML estimators. Therefore, under regularity conditions, they are asymptotically fully efficient, i.e., they are unbiased and minimum variance bound (MVB) estimators.

iii. Even for small sample sizes, they are highly efficient.

iv. They are robust.

It should be noted that weights $\beta_{(j)}$ in (12) have half-umbrella ordering, i.e., they are a decreasing sequence of positive numbers in the direction of the long tail. Therefore, weights $\beta_{(j)}$ given to the extreme residuals deplete the dominant effect of long tail and outliers. This is a very important property for achieving robustness, see for example Tiku & Akkaya (2004). On the other hand, in LS method, all $e_{(j)}$ receive the same weight. This exposes the LS estimators to the dominant effect of long tail and outliers making them nonrobust.

5. Comparison of Estimators

In this section, we compare the ML, MML and LS estimators of the model parameters in terms of means, variances and mean square errors (MSE) for some representative values of the skewness parameter λ. All the simulations are based on $[100,000/n]$ Monte Carlo runs. In the simulation study, we use $a=3,5,n=5,10,15,20$ and $\alpha=0.05$, however, we just reproduce the results for $a=3$ for the sake of brevity. Without loss of generality, we choose the following setting in our simulation: $\mu_i(\mu+\alpha_i)=0(i=1,2,\ldots,1)$ and $\sigma=1$.

Here, it should be noted that we are interested in λ values satisfying the property $0.4<[P(X>E(X))]<0.6$ in the context of experimental design. We, therefore use λ values satisfying the mentioned condition, i.e. we take $-1<\lambda<1$ from now on. Simulation results are given in Table 2.

From Table 2, it is seen that both the ML and the MML estimators are more efficient than the LS estimators of μ_i and σ when the skewness parameter λ is close to 1. When the skewness parameter λ is close to 0 all the three estimators have similar efficiencies as expected. Because, $SN(\lambda)$ distribution reduces to normal distribution when λ is equal to 0; in that case, algebraic forms of the ML and the MML estimators are exactly the same with the corresponding LS estimators of the unknown parameters.

It is interesting to note that relative efficiencies (REs) of the ML and the MML estimators decrease as the sample size n increases.

TABLE 2: Means, variances and MSEs for the LS, ML and MML estimators of μ_i and σ.

	Mean			Variance			MSE			RE	
n	$\hat{\mu}_{i,LS}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$	$\hat{\mu}_{i,LS}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$	$\hat{\mu}_{i,LS}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$
$\lambda = 0$											
5	−0.002	−0.002	−0.002	0.201	0.201	0.201	0.201	0.201	0.201	100	100
10	−0.005	−0.005	−0.005	0.097	0.097	0.097	0.097	0.097	0.097	100	100
15	−0.003	−0.003	−0.003	0.067	0.067	0.067	0.067	0.067	0.067	100	100
20	−0.001	−0.001	−0.001	0.048	0.048	0.048	0.048	0.048	0.048	100	100
$\lambda = 0.4$											
5	0.021	0.008	0.009	0.186	0.186	0.186	0.186	0.186	0.186	100	100
10	0.017	0.005	0.006	0.091	0.089	0.090	0.091	0.089	0.090	98	99
15	0.011	−0.001	0.001	0.063	0.060	0.061	0.063	0.060	0.061	96	97
20	0.015	0.002	0.002	0.048	0.046	0.046	0.048	0.046	0.046	96	96
$\lambda = 0.7$											
5	0.055	0.012	0.013	0.165	0.166	0.166	0.168	0.166	0.166	99	99
10	0.053	0.006	0.008	0.082	0.082	0.082	0.085	0.082	0.082	97	97
15	0.051	0.003	0.004	0.053	0.054	0.054	0.056	0.054	0.054	96	96
20	0.055	0.006	0.008	0.041	0.042	0.042	0.044	0.026	0.042	96	96
$\lambda = 1.0$											
5	0.104	0.023	0.024	0.139	0.139	0.139	0.149	0.142	0.142	95	95
10	0.107	0.015	0.018	0.072	0.072	0.072	0.083	0.073	0.073	88	88
15	0.104	0.011	0.012	0.045	0.045	0.045	0.056	0.046	0.046	82	82
20	0.101	0.004	0.007	0.034	0.034	0.034	0.044	0.034	0.034	77	77
n	$\hat{\sigma}_{i,LS}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$	$\hat{\sigma}_{i,LS}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$	$\hat{\sigma}_{i,LS}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$
$\lambda = 0$											
5	0.981	0.981	0.981	0.040	0.040	0.040	1.003	1.003	1.003	100	100
10	0.992	0.992	0.992	0.019	0.019	0.019	1.002	1.002	1.002	100	100
15	0.994	0.994	0.994	0.012	0.012	0.012	1.000	1.000	1.000	100	100
20	0.996	0.996	0.996	0.009	0.009	0.009	1.002	1.002	1.002	100	100
$\lambda = 0.4$											
5	0.997	0.987	0.987	0.038	0.038	0.038	1.033	1.013	1.013	98	98
10	0.991	0.981	0.981	0.018	0.018	0.018	1.001	0.981	0.981	98	98
15	1.007	0.997	0.997	0.012	0.012	0.012	1.026	1.006	1.006	98	98
20	0.997	0.988	0.989	0.010	0.009	0.010	1.006	0.986	0.986	98	98
$\lambda = 0.7$											
5	1.008	0.981	0.982	0.042	0.039	0.039	1.060	1.001	1.005	95	95
10	1.019	0.991	0.991	0.020	0.019	0.019	1.059	1.005	1.005	95	95
15	1.022	0.994	0.994	0.012	0.012	0.012	1.059	1.001	1.001	95	95
20	1.059	0.997	0.997	0.008	0.008	0.008	1.061	1.003	1.003	95	95
$\lambda = 1.0$											
5	1.024	0.988	0.989	0.049	0.043	0.044	1.097	0.988	1.002	90	91
10	1.044	0.990	0.996	0.020	0.018	0.018	1.111	0.999	1.012	90	91
15	1.045	0.992	0.996	0.013	0.012	0.012	1.106	0.998	1.005	90	91
20	1.048	0.996	0.998	0.010	0.009	0.009	1.109	1.003	1.006	90	91

Robustness: In this study, we use the following definition of robustness. An estimator is called robust if it is fully efficient under the assumed model and maintains high efficiency under the plausible alternatives of the assumed model, (see, Tiku & Akkaya 2004). Assume, for illustration, that the true model in the simulation study is taken to be $SN(0, 1, 1)$. We use the following sample models to represent a large number of plausible alternatives.

Sample Models:

Model (1): Dixon's outlier model: $(n - 1)$ observations come from $SN(0, 1, 1)$ but one observation (we do not know which one) comes from $SN(0, 2, 1)$

Model (2): Dixon's outlier model: $(n - 1)$ observations come from $SN(0, 1, 1)$ but one observation (we do not know which one) comes from $SN(0, 4, 1)$

Model (3): Mixture model: $0.90SN(0, 1, 1) + 0.10SN(0, 1, 0.4)$

Model (4): Contamination model: $0.90SN(0, 1, 1) + 0.10N(0, 1)$.

Given in Table 3 are the simulated values of the means, variances and MSEs for the ML, the MML and the LS estimators of the model parameters $\mu_i(i = 1, 2, \dots, a)$ and σ under the alternative models. We simply reproduce the results for μ_1 since they are all similar. We also give the REs of the ML and the MML estimators with respect to the LS estimators.

Table 3: Means, variances and MSEs for the LS, ML and the MML estimators of μ_i and σ for the alternative models.

	Mean			Variance			MSE			RE	
n	$\hat{\mu}_{i,LS}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$	$\hat{\mu}_{i,LS}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$	$\hat{\mu}_{i,LS}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$	$\hat{\mu}_{i,ML}$	$\hat{\mu}_{i,MML}$
Model 1											
5	0.096	−0.011	0.004	0.206	0.203	0.203	0.216	0.203	0.203	94	94
10	0.101	−0.011	0.007	0.092	0.092	0.092	0.102	0.092	0.092	90	90
15	0.081	−0.022	−0.016	0.059	0.059	0.059	0.066	0.059	0.059	90	90
20	0.089	−0.010	−0.006	0.040	0.040	0.040	0.049	0.041	0.041	84	84
Model 2											
5	0.280	0.102	0.199	0.507	0.540	0.489	0.585	0.550	0.528	94	91
10	0.167	0.023	0.119	0.160	0.166	0.154	0.187	0.167	0.169	89	89
15	0.143	0.017	0.112	0.092	0.095	0.090	0.112	0.095	0.097	84	89
20	0.110	0.010	0.084	0.057	0.057	0.056	0.070	0.058	0.060	82	89
Model 3											
5	0.057	−0.024	−0.024	0.160	0.163	0.163	0.164	0.164	0.164	100	100
10	0.064	−0.029	−0.026	0.079	0.081	0.081	0.084	0.082	0.082	97	97
15	0.065	−0.029	−0.029	0.048	0.049	0.049	0.052	0.050	0.050	96	96
20	0.068	−0.026	−0.026	0.038	0.038	0.038	0.043	0.039	0.039	91	91
Model 4											
5	0.052	−0.034	−0.033	0.148	0.147	0.148	0.147	0.147	0.147	99	99
10	0.055	−0.036	−0.036	0.081	0.080	0.080	0.082	0.080	0.080	98	98
15	0.059	−0.037	−0.037	0.056	0.055	0.055	0.059	0.057	0.057	97	97
20	0.051	−0.031	−0.032	0.039	0.039	0.039	0.041	0.039	0.039	95	95
n	$\hat{\sigma}_{i,LS}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$	$\hat{\sigma}_{i,LS}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$	$\hat{\sigma}_{i,LS}$	$\hat{\sigma}_{i,ML}$	$\hat{\sigma}_{i,MML}$
Model 1											
5	1.313	1.242	1.243	0.107	0.096	0.095	1.830	1.639	1.639	90	90
10	1.201	1.138	1.136	0.043	0.038	0.038	1.485	1.334	1.331	90	90
15	1.153	1.094	1.092	0.023	0.021	0.021	1.352	1.219	1.217	90	90
20	1.129	1.072	1.071	0.016	0.015	0.015	1.292	1.164	1.163	90	90
Model 2											
5	2.178	2.051	2.020	0.547	0.435	0.426	5.292	4.702	4.521	88	86
10	1.738	1.622	1.599	0.228	0.212	0.238	3.251	2.843	2.721	87	84
15	1.565	1.467	1.441	0.145	0.114	0.134	2.595	2.266	2.179	87	84
20	1.455	1.348	1.340	0.096	0.079	0.087	2.214	1.897	1.861	85	84
Model 3											
5	1.046	0.994	1.001	0.045	0.041	0.042	1.141	1.030	1.043	90	91
10	1.063	1.009	1.016	0.023	0.020	0.020	1.154	1.040	1.053	90	91
15	1.071	1.016	1.021	0.015	0.013	0.013	1.161	1.046	1.057	90	91
20	1.067	1.013	1.018	0.011	0.009	0.009	1.149	1.037	1.046	90	91
Model 4											
5	1.074	1.015	1.030	0.056	0.050	0.052	1.212	1.085	1.113	90	92
10	1.081	1.025	1.034	0.027	0.024	0.025	1.196	1.077	1.095	90	92
15	1.088	1.034	1.040	0.016	0.015	0.015	1.201	1.084	1.097	90	91
20	1.087	1.032	1.039	0.012	0.010	0.011	1.194	1.077	1.091	90	91

It can be seen that the ML and the MML estimators are robust owing to the reason mentioned at the end of the Section 4.

6. Hypothesis Testing

In one-way ANOVA, our aim is to compare the equality of treatment effects, in other words, to test the following null hypothesis

$$H_0 : \alpha_i = 0, i = 1, 2, \ldots, a \tag{15}$$

against the alternative hypothesis

$$H_1 : \text{at least one}\, \alpha_i \neq 0.$$

Traditionally, for testing the null hypothesis given in (15) the following test statistics based on the LS estimators are used

$$F_{LS} = \frac{n \sum_{i=1}^{n} \hat{\alpha}_{i,LS}}{(a-1)\hat{\sigma}^2_{LS}} \tag{16}$$

As mentioned earlier, power of F_{LS} is very sensitive to non-normality and to data anomalies. Therefore, in this paper, we propose the following test statistics based on the ML and the MML estimators as an alternative to the test statistic given in (16),

$$F_{ML} = \frac{n\sum_{i=1}^{n}\hat{\alpha}_{i,ML}}{(a-1)(1-\lambda^2t^2)\hat{\sigma}^2_{ML}}, \; F_{MML} = \frac{m\sum_{i=1}^{n}\hat{\alpha}_{i,MML}}{(a-1)\hat{\sigma}^2_{MML}} \tag{17}$$

Large values of F_{ML} and F_{MML} lead to the rejection of H_0. For large n values, distribution of both F_{ML} and F_{MML} are central F with degrees of freedom $(a-1, N-a)$. On the other hand, for small n values, we use Monte Carlo simulation study to verify how close their null distribution is to central F. We simulate the Type I errors of F_{ML} and F_{MML} by computing the following probabilities

$$P(F_{ML} \geq F_\alpha(a-1, N-a)|H_0) \quad \text{and} \quad P(F_{MML} \geq F_\alpha(a-1, N-a)|H_0), \tag{18}$$

respectively. Table 4 shows that central F distribution with $a-1$ and $N-a$ degrees of freedom provides accurate approximations to the distributions of F_{ML} and F_{MML} even for small n values.

Table 4: Simulated Type I Errors of F_{LS}, F_{ML} and F_{MML} tests $a = 3$; $\alpha = 0.050$.

λ	n	5	10	15	20
	F_{LS}	0.050	0.049	0.048	0.050
0	F_{ML}	0.054	0.050	0.053	0.052
	F_{MML}	0.053	0.051	0.056	0.054
	F_{LS}	0.046	0.055	0.055	0.045
0.4	F_{ML}	0.049	0.054	0.056	0.049
	F_{MML}	0.048	0.055	0.054	0.047
	F_{LS}	0.049	0.054	0.049	0.053
0.7	F_{ML}	0.050	0.052	0.049	0.049
	F_{MML}	0.051	0.054	0.047	0.048
	F_{LS}	0.054	0.048	0.051	0.049
1.0	F_{ML}	0.055	0.049	0.052	0.054
	F_{MML}	0.055	0.049	0.053	0.053

We now compare the power of the F_{ML} and F_{MML} tests with the traditional F_{LS} test by simulating the probabilities

$$P(F_{ML} \geq F_\alpha(a-1, N-a)|H_1) \quad \text{and} \quad P(F_{MML} \geq F_\alpha(a-1, N-a)|H_1), \tag{19}$$

for some representative values of λ. It should be noted that all the observations are divided by their standard errors. A constant d is added to the observations in the first and third treatments and a constant $2d$ is subtracted from the observations in the second treatment. Simulation results showing the power comparisons of the proposed tests with the LS based test are given in Table 5.

From Table 5 it is clear that power of F_{LS} , F_{ML} and F_{MML} are very similar when λ is close to 0. When λ approaches 1, F_{ML} and F_{MML} seem more powerful than the F_{LS}, but the differences are not very attractive. This is not surprising due to the fact that the quadratic form of a skew-normal distributed random variable has the chi-square distribution (Azzalini 1985, Gupta & Huang 2002).

Table 5: Power values of the F_{LS}, F_{ML} and F_{MML} tests: $a = 3$,$n = 10$; $\alpha = 0.050$.

λ	0			0.4			0.7			1.0		
d	F_{LS}	F_{ML}	F_{MML}	F_{LS}	F_{ML}	F_{MML}	F_{LS}	F_{ML}	F_{MML}	F_{LS}	F_{ML}	F_{MML}
0	0.050	0.050	0.050	0.049	0.049	0.049	0.055	0.056	0.054	0.051	0.052	0.053
0.1	0.09	0.09	0.09	0.09	0.09	0.09	0.08	0.09	0.08	0.08	0.09	0.09
0.2	0.24	0.24	0.24	0.22	0.22	0.22	0.20	0.21	0.20	0.17	0.19	0.19
0.3	0.48	0.48	0.48	0.46	0.46	0.46	0.40	0.41	0.40	0.35	0.38	0.37
0.4	0.72	0.72	0.72	0.71	0.71	0.71	0.65	0.66	0.66	0.58	0.61	0.59
0.5	0.90	0.90	0.90	0.89	0.89	0.89	0.84	0.85	0.84	0.78	0.80	0.79
0.6	0.98	0.98	0.98	0.97	0.97	0.97	0.95	0.96	0.96	0.91	0.92	0.92
0.7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.97	0.98	0.97

Robustness: We use the following definitions of robustness formulated by Box (1953). See also Box & Tiao (1964), Tiku, Tan & Balakrishnan (1986).

Criterion robustness: If the Type I error of a test is not substantially higher under plausible alternatives than that attained under an assumed model, the test is said to have criterion robustness.

Efficiency robustness: If the power of a test is the highest possible (or nearly so) under an assumed model but stays high for all plausible models, the test is said to have efficiency robustness.

In this section, our aim is to identify the affect of deviations from an assumed model on the Type I error and the power of the proposed tests. For this purpose, we use the sample models given in Section 5. These simulated values of the power of the proposed tests and the F_{LS} test are given in Table 6.

Table 6: Values of the power for the alternatives to $SN(0, 1, 1)$: $a = 3$, $n = 10$; $\alpha = 0.050$.

	(1)			(2)			(3)			(4)		
d	F_{LS}	F_{ML}	F_{MML}	F_{LS}	F_{ML}	F_{MML}	F_{LS}	F_{ML}	F_{MML}	F_{LS}	F_{ML}	F_{MML}
0	0.050	0.053	0.054	0.031	0.055	0.053	0.050	0.052	0.051	0.048	0.054	0.053
0.1	0.06	0.08	0.08	0.06	0.09	0.08	0.08	0.09	0.10	0.09	0.10	0.10
0.2	0.15	0.19	0.19	0.18	0.26	0.27	0.18	0.21	0.22	0.18	0.21	0.21
0.3	0.28	0.33	0.33	0.35	0.44	0.45	0.35	0.38	0.39	0.35	0.38	0.38
0.4	0.47	0.52	0.52	0.44	0.54	0.55	0.57	0.61	0.61	0.57	0.62	0.63
0.5	0.65	0.70	0.69	0.57	0.67	0.68	0.75	0.78	0.79	0.78	0.81	0.82
0.6	0.81	0.85	0.84	0.70	0.79	0.80	0.90	0.92	0.93	0.91	0.94	0.94
0.7	0.92	0.94	0.95	0.89	0.96	0.96	0.98	0.99	0.99	0.96	0.98	0.98

It is clear from Table 6 that the power of the F_{ML} and F_{MML} tests are much higher than the corresponding F_{LS} test for all the sample models, i.e., Model (1) through Model (4). For $d = 0$, the values represent Type I errors. Then it is said that proposed tests have criterion robustness as well as the efficiency robustness.

7. Application

Consider the data given in Montgomery (2005); pertaining to the relationship between the radio frequency power setting and the etch rate for plasma. This is an example of a one-way ANOVA with 4 levels of the factor and 5 replicates. The data is given in Table 7.

Table 7: Radio Frequency Data.

160 W	180 W	200 W	220 W
575	565	600	725
542	593	651	700
530	590	610	715
539	579	637	685
570	610	629	710

To identify the distribution of the error terms, we use the Q-Q plot technique, one of the well-known and widely used graphical techniques. The Q-Q plot of normal distribution is shown in Figure 1. On the other hand, among the Q-Q plots of the residuals obtained for various different values of the skewness parameter λ, $SN(\mu, \sigma, \lambda = 1)$ adequately models the residuals, since the observations do not deviate very much from the straight line, see Figure 2.

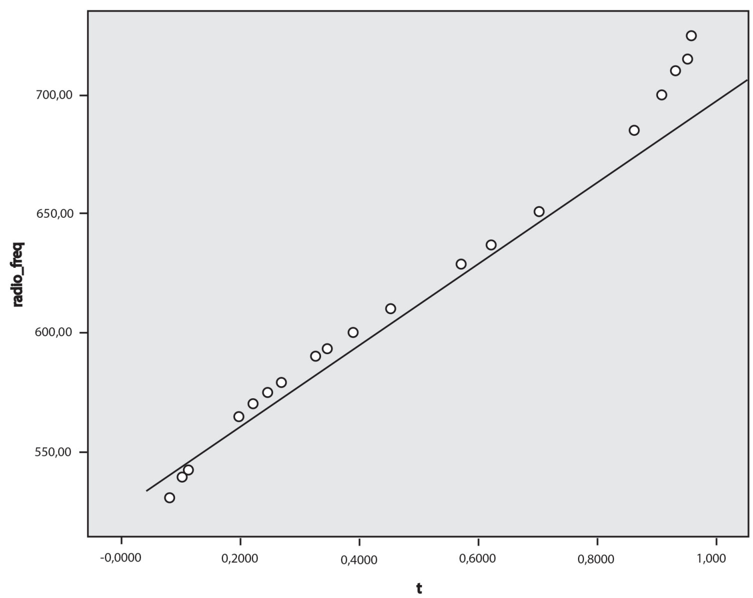

Figure 1: Q-Q plot of the residuals for normal distribution.

When we take the skewness parameter λ as 1, parameter estimates and calculated F values are obtained as shown in Table 8).

The ML and the MML estimates of μ_i are very close to the LS estimate of μ_i with smaller standard errors. All the three tests are consistent in rejecting the null hypothesis, H_0: there is no difference between the radio frequency powers.

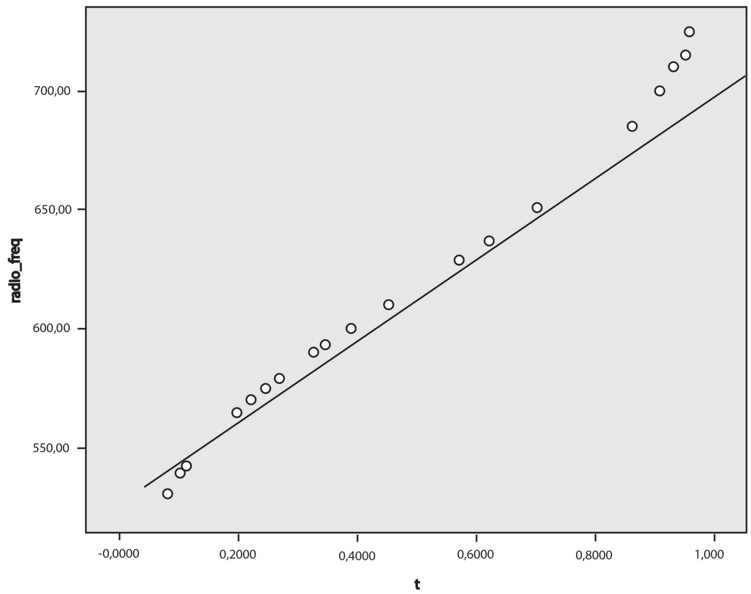

Figure 2: Q-Q plot of the residuals for $SN(\lambda = 1)$.

Table 8: The parameter estimates and the calculated F values.

	μ	α_1	α_2	α_3	α_4	σ	F
LS	617.75	−66.55	−30.35	7.65	89.25	22.125	66.797*
ML	616.28	−64.22	−34.12	8.28	90.08	19.818	84.041*
MML	616.34	−64.05	−34.68	8.08	90.65	21.108	71.125*

*Reject H_0

However, the p values for F_{ML} and F_{MML} are much smaller than the p value of the F_{LS}. This is due to the smaller standard errors of the ML and the MML estimators. Therefore, their results are more reliable than normal theory results.

8. Conclusion

Traditionally, LS estimators and the tests based on them are used in the context of experimental design. However, efficiencies of the LS estimators are low when the usual normality assumption is not satisfied. They are also not robust to departures from normality.

In this paper, we derived estimators of the model parameters in one-way ANOVA by using the ML and the MML methodologies. New test statistics based on these estimators were proposed for testing the equality of the treatment effects when the distribution of the error terms is skew-normal. $SN(\lambda)$ distribution covers the normal and normal-like distributions with different skewness and kurtosis values. Therefore, it provides very flexible and simple alternative model for the normal distribution in most practical problems.

Simulation studies show that the ML and the MML estimators and the tests based on them are more efficient and robust than the corresponding LS versions thereof.

It can also be seen that there is no significant difference between the methodologies based on ML and MML even for small sample sizes. The methodology based on ML is somewhat preferable than the methodology based on MML in terms of efficiency and power. On the other hand, the methodology based on MML is computationally feasible and less time consuming.

Acknowledgements

The authors would like to thank the referees for their many helpful comments and suggestions.

[Received: November 2013 — Accepted: June 2014]

References

Acitas, S., Kasap, S., Senoglu, B. & Arslan, O. (2013), 'Robust estimation with the skew t_2 distribution', *Pakistan Journal of Statistics* **29**(4), 409–430.

Arrellano-Valle, R., Bolfarine, H. & Lachos, V. (2005), 'Skew-normal linear mixed models', *Journal of Data Science* **3**, 415–438.

Arslan, O. & Genc, A. (2009), 'The skew generalized t (sgt) distribution as the scale mixture of a skew exponential distribution and its application in robust estimation', *Statistics* **43**, 481–498.

Azzalini, A. (1985), 'A class of distributions which includes the normal ones', *Scandinavian Journal of Statistics* **12**, 171–178.

Azzalini, A. (1986), 'Further results on a class of distributions which includes the normal ones', *Statistica* **46**, 199–208.

Azzalini, A. (2005), 'The skew-normal distribution and related multivariate families (with discussion)', *Scandinavian Journal of Statistics* **32**, 159–188.

Bowman, K. & Shenton, L. (2001), 'Weibull distributions when the shape parameter is defined', *Computational Statistics and Data Analysis* **36**, 299–310.

Box, G. (1953), 'Non-normality and test of variances', *Biometrika* **40**, 336–346.

Box, G. & Tiao, G. (1964), 'A Bayesian approach to the importance of assumptions applied to the comparison of variances', *Biometrika* **51**, 153–167.

Donaldson, T. (1968), 'Robustness of the f-test to errors of both kinds and the correlation between the numerator and denominator of the F ratio', *Journal of American Statistical Association* **63**(322), 660– 667.

Garay, A., Lachos, V. & Abanto-Valle, C. (2011), 'Nonlinear regression models based on scale mixtures of skew-normal distributions', *Journal of Korean Statistical Society* **40**, 115–124.

Garay, A., Lachos, V., Labra, F. & Ortega, E. (2013), 'Statistical diagnostics for nonlinear regression models based on scale mixtures of skew normal distributions', *Journal of Statistical Computation and Simulation* **84**, 1761–1778.

Geary, R. (1947), 'Testing for normality', *Biometrika* **34**, 209–242.

Gupta, A. & Huang, W. (2002), 'Quadratic forms in skew normal variates', *Journal of Mathematical Analysis and Applications* **273**, 558–564.

Huber, P. (1981), *Robust Statistics*, Jonh Wiley, New York.

Islam, M. & Tiku, M. (2004), 'Multiple linear regression model under nonnormality', *Communication Statistics -Theory Methods* **33**, 2443–2467.

Kantar, Y. & Senoglu, B. (2008), 'A comparative study for the location and scale parameters of the weibull distribution with given shape parameter', *Computers and Geosciences* **34**, 1900–1909.

Lachos, V., Bandyopadhyay, D. & Garay, A. (2011), 'Heteroscedastic non linear regression models based on scale mixtures of skew-normal distributions', *Statistics and Probability Letters* **81**, 1208–1217.

Lachos, V., Bolfarine, H., Arellano-Valle, R. B. & Montenegro, L. (2007), 'Likelihood based inference for multivariate skew normal regression models', *Communications in Statistics* **36**(9), 1769–1786.

Lachos, V., Ghosh, P. & Arellano-Valle, R. (2010), 'Likelihood based inference for skew-normal independent linear mixed models', *Statistica Sinica* **20**, 303–322.

Martínez-Flórez, G., Vergara-Cardozo, S. & González, L. M. (2013), 'The family of Log-Skew-Normal Alpha-power distributions using precipitation data', *Revista Colombiana de Estadística* **36**(1), 43–57.

Montgomery, D. (2005), *Design and Analysis of Experiments*, John Wiley & Sons Inc., United States of America.

Pearson, E. (1932), 'The analysis of variance in cases of nonnormal variation', *Biometrika* **23**, 114–133.

Pereira, J. R., Marques, L. A. & da Costa, J. M. (2012), 'An empirical comparison of EM initialization methods and model choice criteria for mixtures of Skew-Normal distributions', *Revista Colombiana de Estadística* **35**(3), 457–478.

Senoglu, B. & Tiku, M. (2001), 'Analysis of variance in experimental design with nonnormal error distributions', *Communication Statistical Theory Methods* **30**, 1335–1352.

Senoglu, B. & Tiku, M. (2002), 'Linear contrasts in experimental design with non-identical error distributions', *Biometrical Journal* **44**(3), 359–374.

Spjotvoll, E. & Aastveit, H. (1980), 'Comparison of robust estimators on some data from field experiments', *Scandinavian Journal of Statistics* **7**, 1–13.

Srivastava, A. (1959), 'Effect of nonnormality on the power of the analysis of variance test', *Biometrika* **46**, 114–122.

Tan, W. & Tiku, M. (1999), *Sampling Distributions in Terms of Laguerre Polynomials with Applications*, New Age International (formerly, Wiley Eastern), New Delhi.

Tiku, M. (1967), 'Estimating the mean and standard deviation from censored normal samples', *Biometrika* **54**, 155–165.

Tiku, M. & Akkaya, A. (2004), *Robust Estimation and Hypothesis Testing*, New Age International, New Delhi.

Tiku, M., Tan, W. & Balakrishnan, N. (1986), *Robust Inference*, Marcel Dekker, New York.

Tukey, J. (1960), A survey of sampling from contaminated distributions, *in* I. Olkin, ed., 'Contributions to Probability and Statistics', Stanford University Press, Stanford.

Xie, F., Wei, B. & Lin, J. (2009), 'Homogeneity diagnostics for skew-normal nonlinear regression models', *Statistics and Probability Letters* **79**, 821–827.