

Efficient Rank-Based Analysis of Multilevel Models for the Family of Skew-t Errors

Sehar Saleem^{1*}, Rehan Ahmad Khan Sherwani¹

* Corresponding Author



1. College of Statistical and Actuarial Sciences, University of the Punjab, Lahore 5400- Pakistan, seharsaleem87@gmail.com, rehan.stat@pu.edu.pk

Abstract

Rank-based analysis of linear models is based on selecting an appropriate score function. The information about the shape of the underlying distribution is necessary for the optimal selection; leading towards asymptotically efficient analysis. In this study, we analyzed the multilevel model with cluster-correlated error terms following a family of skew-t distribution with the rank-based approach based on score function derived for the class of skew-normal distribution. The rank fit is compared with the Restricted Maximum Likelihood (REML) estimation in terms of validity and efficiency for different sample sizes. A Monte Carlo simulation study is carried out over skewed-t and contaminated-t distribution with a range of skewness parameter from moderately to highly skewed. The standard error of regression coefficients is significantly reduced in the rank-based approach and further reduces for a large sample size. Rank-based fit appeared asymptotically efficient than REML for each shape parameter of skewness in skew-t and contaminated-t distribution computed through a calculation of precision. The empirical validity of fixed effects is obtained up to the nominal level 0.95 in REML but not rank-based with skew-normal score function.

Key Words: Multilevel Models, Rank-based, REML, Skew-t, Skew-normal

1. Introduction

Several experimental and observational studies often occur resulting in cluster-correlated data, for example in clinical designs with multiple centers known as clusters, repeated measurement designs where individuals are supposed as clusters. In the randomized complete block, design blocking is a random effect, and blocks serve the same as clusters, both terms can be used interchangeably (Auda et al., 2017). One example of these mixed effect models is a multilevel model for incorporating cluster-correlated errors. The observations between clusters are independent and within each cluster dependent. Hence the assumption of the independent error term in linear models is violated. The most common conventional technique to estimate the mixed effect model is Restricted Maximum Likelihood (REML) and methods based on the ordinary least square fitting (OLS). Inferential procedures like testing of hypothesis, confidence intervals, standard error of fixed effects are affected if the underlying likelihood is non-normal due to the presence of outliers (Auda et al., 2017). Neglecting the dependent structure in the model causes underestimating the standard error of the regression coefficient and inflating type-I error (Goldstein, 1995). The only small number of contaminated values can spoil the REML fit leading to biased estimates. The robust estimation methods are least sensitive to outliers and therefore provides an attractive alternative to REML and OLS for linear models. One of the commonly used robust analysis is rank-based fitting of linear models. Auda *et al.*, performed a detailed comparison of the rank-based method with traditional estimation techniques REML of the mixed model. They discussed the power and efficiency of rank-based methods over normal and some non-normal situations (Auda et al., 2017). Kloke, McKean, & Rashid (2009), has extended this rank theory for linear models with cluster-correlated errors. This rank theory provides a comprehensive methodology including general linear testing of hypothesis, confidence intervals, and robust diagnostic measures for model fit (Hettmansperger & McKean, 2010). It replaces the Euclidean norm in the least square fitting with the pseudo norm that produces robust fixed effect estimates (Kloke & McKean, 2014). The rank-based theory is

generally efficient however, this efficiency can be optimized depending on the distribution of random errors (McKean & Kloeke, 2014). If the form of the underlying distribution is known, the corresponding selection of the rank-based procedure would provide optimal analysis. This study is focused on rank-based analysis with score function that is suitable for skew-normal (SN) class of distributions of the error term in linear models. This study is focused only on the case of skew-t errors and showed that enhanced efficiency in rank-based method is attained by using suitable score function and the method is far more powerful and efficient as compared to REML. While its application is presented in detail for the skew-t (ST) class of distribution by discussing its validity and efficiency over several different values of the skewness parameter. Skew-normal is a rich class of distribution presented by (Azzalini, 1985). This family of distribution has a shape parameter α , $-\infty < \alpha < \infty$ to control the skewness of each distribution in this class. Because the distribution is symmetric in this class, it can be left or right-skewed depending upon the -ve and +ve value of α respectively. The skew-normal distribution reduces to simple normal-distribution when the shape parameter $\alpha = 0$. This family of distribution has been used in numerical studies by (Adelchi Azzalini & Capitanio, 2003), (Azzalini & Genton, 2008) and (Azzalini & Capitanio, 2014). This study is intended to facilitate valid and precise inferences for skewed data occurred in various practical situations. Some common examples of positive skewness might include income and expenditure of families, prices, number of accidents claimed by insurance companies, number of children etc. are the reasonable events to achieve low boundary usually at zero and extremely large values are also possible to occur. Negative skewness usually less often occurs in real problems, but some situations like students' scores on easy tests for which students are especially motivated and another example includes results of Olympic long jump by athletes. Also retirement ages may have negative skewness: mostly people retire after the age of 60 or 65 but some professions give retirement in early ages, similarly in developed countries age at death is negatively skewed. One of the popular distributions to handle these kind of data sets is skewed-t distribution.

The student's t-distribution is unable to capture the asymmetry however, it efficiently captures values away from mean and heavy tails (Basalamah, 2017). Several methodologies in literature are proposed to incorporate asymmetry. We considered a widely held approach proposed (Arellano-Valle & Azzalini, 2013). Skew-t distribution from this skew-symmetric rich class of distributions is considered for the generation of random errors. Likewise, the skew-normal distribution, skewness of the skew-t distribution is also controlled by a shape parameter α , $-\infty < \alpha < \infty$. The appropriate score function for skew-normal error terms is developed in package `sn` (Azzalini, 2014). This package contains suitable functions to generate and manipulate skew-normal and skew-t random variants. This package can be downloaded at <https://CRAN.R-project.org/package=sn>. The detail about data examples and functions used in this package are available at <http://azzalini.stat.unipd.it/SN>. We applied the score function suitable for skew-normal error distribution from this package and skew-t error terms are generated from multilevel models. The rank-based fit is obtained by using R package `jrfit` and REML estimates are achieved through package `lme4`. We performed a Monte Carlo simulation study to discuss the robustness and efficiency of rank-based procedures developed for skew-normal and applied over a class of skew-t distributions at different sample sizes.

In this study, we conducted a rank-based analysis for a skew-t family of distribution based on skew-normal score function. A simulation study is performed with the rank-based procedure based on a shape parameter α of the skew-normal score function. The analysis for correct α is fully efficient. The rank fit also shows quite similar results with neighborhood values of the correct α . The effect of sample size in the multilevel model is also considered because a larger sample size reduces standard error (SE) leading towards efficient analysis. Thus the analysis is repeated over three different sample sizes small to moderate and the rank-based fit is compared with the REML approach. The findings of this study confirm the outclass performance of skew-normal procedure over the traditional approach REML in the neighborhood of the correct α applied on skew-t level-1 and level-2 random errors in the multilevel model.

2. Representation of rank-based analysis

2.1 General Linear Model

Al-Shomrani presented the linear regression model as shown (Al-shomrani, 2003):

$$Y = \alpha 1 + X\beta + e. \quad (1)$$

Where Y is the dependent variable, X is $n \times p$ the design matrix, β is the $p \times 1$ vector of slope parameters, α is the intercept, and e is a $n \times 1$ vector of residuals that are *i.i.d*. The density function of the error term is denoted as f and distribution function is denoted as F . The design matrix is assumed to have a full column rank and centered as the model includes the intercept. Rank-based estimates are invariant to the intercept.

2.2 Random Intercept Model

The general form of the random intercept multilevel model in combined form is presented (Goldstein, 1995):

$$y_{ij} = \gamma_{00} + \beta_1 x_{ij} + e_{ij} + u_{0j}. \quad (2)$$

Where y_{ij} is a dependent variable measured at level-1 (i subscript refers to individual-level variation and j refers to group-level variation), x_{ij} is an independent variable measured at level-1, β_1 is the fixed slope for each group. e_{ij} is a random error obtained from the level-1 regression equation. γ_{00} is the overall intercept, it is a fixed effect i.e., invariant over the clusters. It shows a unique effect of cluster j on the intercept. Also u_{0j} is a random error that shows the deviation of group-level slopes from the overall slope. It is assumed to be normally and independently distributed across individuals with distribution function F and density function f .

2.3 Brief Overview of Rank Theory

In rank-based theory, a pseudo norm is used instead of the Euclidean norm in OLS. Thus the robust estimate of $\hat{\beta}_\varphi$ can be defined as (John D Klope et al., 2009):

$$\hat{\beta}_\varphi = \arg \min \|Y - X\beta\|_\varphi. \quad (3)$$

This joint rank-based estimator of β is called Jaeckel's dispersion function (Jaeckel, 1972). This is a convex function of β as it is defined in terms of the norm. It is efficient and robust in Y -space (J. D. Klope & McKean, 2012). Al-Shomrani (2003) defined this pseudo norm by

$$\|Y - X\beta\|_\varphi = \sum_{i=1}^N a \left[R(y_{ij} - x'_{ki}\beta) \right] (y_{ij} - x'_{ki}\beta). \quad (4)$$

Where R denotes the rank of v_i among v_1, \dots, v_n and these are invariant to constant shifts.

y_{ij} = Dependent variable for the individual i in a cluster j

x'_{ki} = Corresponding $p \times 1$ vector of covariates

A suitable selection of score function according to the shape of error distribution is required for optimal analysis. The score function $\varphi(u)$ is a non-decreasing square-integrable function bounded in the interval $(0,1)$. The score function can be standardized such that $\int \varphi(u)du = 0$ and $\int \varphi^2(u)du = 1$. Scores are calculated as $a[t] = \varphi\left[\frac{t}{N+1}\right]$ and these

scores sum equal to zero. Some score functions make assumptions about the distribution of the raw scores and few of them do not make any assumptions such as the Wilcoxon score function. If the underlying distribution of the error term is known, Hájek & Sidák, (1967) showed that optimal scores can be found as:

$$\varphi_f(u) = \varphi(u) = \frac{-f'(F^{-1}(u))}{f(F^{-1}(u))}. \quad (5)$$

Where $f(u)$ and $F(u)$ are probability density function (pdf) and cumulative density function (cdf) respectively of the respective error distribution. In the vast set of skew-symmetric distributions, a remarkable case is a skew-normal distribution. It is a superset of a normal family with an additional parameter to control skewness (Arellano-Valle & Azzalini, 2013). Its density function in the univariate case is defined as:

$$\frac{2}{\omega} \varphi(z) \Phi(\alpha z), \quad z = \frac{x - \xi}{\omega}. \quad (6)$$

McKean & Klope (2014) considered the location and scale parameter 0 and 1 respectively, leading to expression of the pdf

$$f(x) = 2\phi(x)\Phi(\alpha x). \quad (7)$$

Where the domain of parameter α is $-\infty < \alpha < \infty$. The pdf and cdf of the standard normal distribution are denoted as $\phi(x)$ and $\Phi(x)$ respectively. Within the context of skew-symmetric distributions, Arellano-Valle & Azzalini (2013) presented a skew-t (ST) family of distribution with pdf at x ($x \in \mathbb{R}$):

$$\frac{2}{\omega} t(z; \nu) T\left(\alpha z \sqrt{\frac{\nu+1}{\nu+z^2}}; \nu+1\right), \quad z = \frac{x-\xi}{\omega} \quad (8)$$

Where $t(,; \tau)$ is the density function and $T(,; \tau)$ denotes the distribution function of the student's t distribution with τ degree of freedom (df). For a standard t version of this pdf includes location $\xi = 0$ and scale parameter $\omega = 1$. Then the pdf of standard skew-t becomes:

$$2t(x; \nu) T\left(\alpha x \sqrt{\frac{\nu+1}{\nu+x^2}}; \nu+1\right). \quad (9)$$

Azzalini (2013) presented some interesting properties of skew-t distribution and discussed its moments as well. Basalamah (2017) summarized some basic properties as:

Let $X \sim ST(\mu, \sigma, \lambda, \gamma)$,

- If $\mu = 0$ and $\sigma = 0$ then $X \sim st_{\gamma}(\lambda)$
- If df $\gamma \rightarrow \pm\infty$ then $X \sim SN(\mu, \sigma, \lambda)$ and
- If $\lambda = 0$ then $X \sim t_{\gamma}(\mu, \sigma)$.
- If $\lambda \rightarrow \pm\infty$ then $X \sim |t_{\gamma}|$
- The measure of skewness of X is the whole real line when df $\gamma > 3$.

McKean & Kloeke (2014) discussed skew-normal distribution and declared that only positive values of α are sufficient because $X \sim SN(\alpha)$ then $-X \sim SN(-\alpha)$. A similar result for skew-t is defined above in point 4. Parallel to the fact in a normal distribution, positive values of α making the distribution right-skewed and negative values of α forms the distribution left-skewed. When $\lambda \rightarrow \infty$ Skew-t density approaches to half t density (Basalamah, 2017). In simulation settings values of α in the interval of -12 and +12 are considered, by using the fact of little difference between skew-normal and half normal distribution when α is near ± 8 (McKean & Kloeke, 2014). Figure 2 shows the graphs of the skew-t pdf's and their corresponding skew-normal scores for the shape parameter $\alpha = -5, 0, 6$. In the left panel of figure 2, the pdf of $\alpha = -5$ is left-skewed and $\alpha = 6$ is right-skewed. It is obvious the pdf for $\alpha = 0$ is symmetric. The score function corresponding to right tail pdf accentuates left tails relatively high. Similarly, score function for left-skewed pdf, put more weight on right tails than left tails.

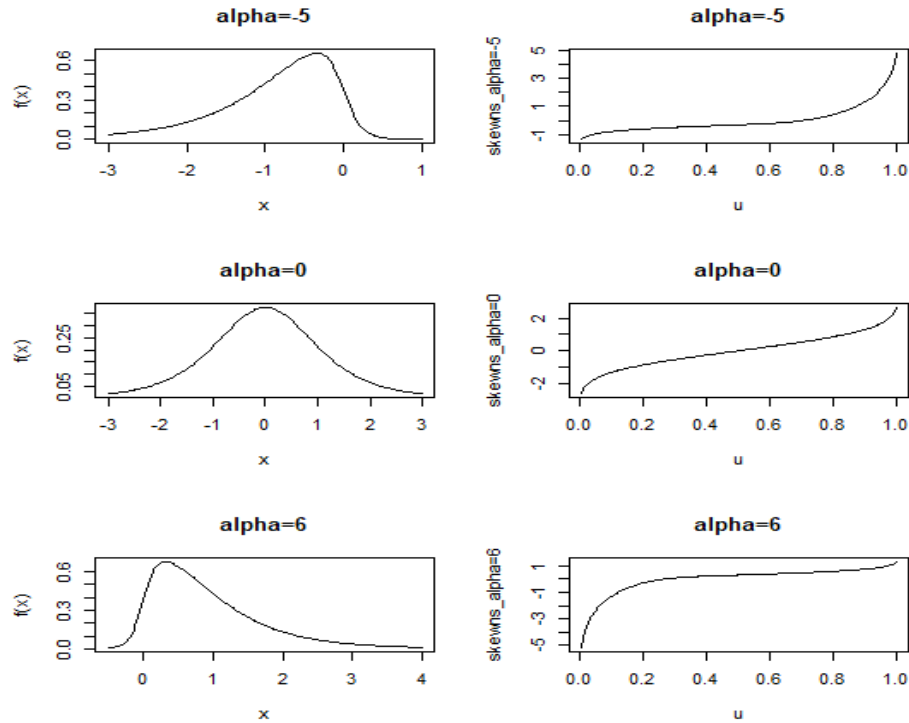


Figure 1: Plot of PDF's of skew-t distribution with shape parameter= -5, 0 and 6 respectively and corresponding skew-normal score functions

3. Simulation Settings

The simulation study is aimed at evaluating rank-based procedures developed for skew-normal distribution applied to the skew-t distribution of level-1 and level-2 error distributions. The score functions are evaluated for multilevel models at different sample sizes. The simulated multilevel model is as follows (McKean & Kloeke, 2014):

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \theta c_i + u_{0j} + e_{ij} \quad j = 1, \dots, k ; i = 1, \dots, n_j \quad (10)$$

Where y_{ij} is the response variable and x_{ij} is a random variable normally distributed $x_{ij} \sim N(0,1)$ measured at level-1. u_{0j} and e_{ij} are random errors at the group level and individual level respectively. Both the random error terms are generated by selected distributions. c_i is an indicator variable of treatment effect where 1 indicates presence and 0 indicates absence. There are three variations of level-1 sample size as: 5, 10, 20 within each level-2 unit, and 5 level-2 units are considered. Two kinds of error distributions are reflected in this simulation study: skew-t distribution and contaminated skew-t distribution. In skew-t distribution both level-1 and level-2 error terms, the degree of freedom (df) is fixed at $\nu = 4$, and shape parameter $\alpha = 7$ in each run. Location and scale parameters are set at 0 and 1 respectively. The fixed effect parameters β_0 , β_1 and θ are set at 0 without loss of generality. The contaminated skew-t model is given below:

$$e_i = (1 - I_{\varepsilon,i})W_i + I_{\varepsilon,i}V_i \quad (11)$$

Where $I_{\varepsilon,i}$ is a random variable with a binomial distribution $(1, \varepsilon = 0.15)$. W is the skew-t random error with shape parameter $\alpha = 7$ and df is $\nu = 4$. V_i is a random variable from t-distribution $V_i \sim t(4)$. A similar setting is used for group-level error too. For robust analysis through rank theory, the score $\varphi_\gamma(u)$ is selected suitable for $\alpha = 7$ in standard skew-normal distribution. Three values of α in the neighborhood of $\alpha = 7$ are considered to show a

comparative analysis of optimal score function i.e. $\alpha = 4, 5, 6, 7, 8, 9, 10$. To investigate the effect of high skewness we also consider $\alpha = 12$.

Algorithm of the simulation study

1. Simulate random variable $X \sim N(0, 1)$, C as a treatment indicator and a block effect.
2. Random error at group level and individual level are generated with standard skew-t distribution with $\text{df } \gamma = 4$ and $\alpha = 4, 5, 6, 7, 8, 9, 10$ and 12.
3. The group size 5 is considered, within each group level-1 sample size is taken 5, 10, and 20. Thus total $n = 25, 50, 100$ is taken for 1000 simulation runs.
4. The rank-based procedure with score function $\varphi_\gamma(u)$ and REML method is applied for all values of α for skew-t (Eq. 10) and contaminated skew-t distribution (Eq. 11) of the error term in the multilevel model.
5. SE of fixed effect β & θ and validity in terms of 95% confidence interval is computed by REML and all rank-based procedures for skew-t and contaminated skew-t distribution.
6. Precision $(\hat{\sigma} / \hat{\tau}_\alpha)^2$ is calculated for the overall comparison of the efficiency of every rank-based method with corresponding REML analysis.

4. Results and Discussion

Table-1 summarizes the results of simulations, comprising rank-based and REML analysis of skewed-t distribution. SE and validity based on a 95% confidence interval $\hat{\beta} \pm t_{\alpha/2, v} SE(\hat{\beta})$ are calculated for two regression coefficients β and θ in the model (Eq. 10) for $n = 25, 50, 100$. All the values of SE of β and θ reducing significantly in the rank-based method as compared to REML, which justifies the robustness of rank theory. With the increase in sample size, the SE decreases a little bit that confirms the effect of a large sample size. The value of precision is calculated greater than 1 indicates the rank-based fit is more precise than REML. And the value of precision less than 1 means REML is better. For each value of the shape parameter, precision is always greater than 1. The most efficient analysis of fixed effects comes out in rank-based analysis within the skew-t distribution depicted from high values of precision 3.61 comparing for $\alpha = 4, 5, 6, 7, 8, 9, 10$. However, quite similar results are obtained for $\alpha = 8$. Regarding validity, an increase in sample size does not endorse to increase in validity. Too small sample sizes may not produce a valid confidence interval. For all procedures, empirical efficiencies should close to the nominal confidence level 0.95. The highest value of precision is obtained when errors are highly right-skewed.

Table 1: Results of rank-based procedures compared with REML for skew-t distribution with shape parameter $\alpha=7$

		Skew-t errors				REML				
Proc	n	β		θ		β		θ		Precision
		SE	Conf	SE	Conf	SE	Conf	SE	Conf	
$\alpha=4$	25	0.39	0.65	0.71	0.12	0.90	0.85	1.61	0.83	2.03
	50	0.32	0.94	0.58	0.94	0.82	0.85	1.48	0.82	2.53
	100	0.30	0.65	0.55	0.82	0.74	0.83	1.35	0.83	2.35
$\alpha=5$	25	0.38	0.69	0.71	0.90	0.88	0.86	1.62	0.86	2.55
	50	0.33	0.36	0.59	0.79	0.82	0.83	1.46	0.83	2.86
	100	0.29	0.34	0.53	0.66	0.81	0.82	1.48	0.81	2.82
$\alpha=6$	25	0.38	0.51	0.67	0.86	0.89	0.88	1.57	0.85	2.48
	50	0.32	0.66	0.56	0.83	0.79	0.84	1.39	0.84	2.98
	100	0.28	0.59	0.52	0.81	0.77	0.84	1.41	0.83	2.93
$\alpha=7$	25	0.38	0.75	0.67	0.86	0.88	0.86	1.55	0.83	2.33
	50	0.30	0.46	0.54	0.66	0.78	0.83	1.43	0.83	3.58
	100	0.28	0.61	0.51	0.90	0.77	0.83	1.38	0.84	3.61
$\alpha=8$	25	0.38	0.43	0.66	0.82	0.89	0.87	1.53	0.85	2.35
	50	0.32	0.81	0.57	0.96	0.83	0.84	1.47	0.83	3.33
	100	0.22	0.27	0.50	0.67	0.82	0.84	1.46	0.83	3.35
$\alpha=9$	25	0.37	0.67	0.66	0.92	0.85	0.86	1.53	0.87	2.62
	50	0.33	0.63	0.56	0.90	0.87	0.84	1.47	0.82	3.17
	100	0.25	0.22	0.46	0.33	0.73	0.83	1.34	0.82	3.57
$\alpha=10$	25	0.39	0.89	0.68	0.93	0.91	0.84	1.62	0.84	1.53
	50	0.31	0.62	0.55	0.85	0.84	0.85	1.47	0.83	1.26
	100	0.26	0.70	0.47	0.86	0.70	0.84	1.27	0.83	1.14
$\alpha=12$	25	0.37	0.66	0.65	0.72	1.31	0.87	2.31	0.82	8.55
	50	0.27	0.75	0.48	0.84	1.16	0.85	2.09	0.83	11.95
	100	0.22	0.61	0.40	0.91	1.03	0.84	1.96	0.83	28.54

Because REML performs worse in such a high skewness and rank-based method still performs best. Surprisingly, the rank-based procedure with score function $\varphi_7(u)$ does not produce nominal confidence close to 0.95. Table-2 comprises a parallel analysis for contaminated skewed-t error distribution of error term. Empirical efficiencies are greatly efficient than REML.

Table 2: Results of rank-based procedures compared with REML for contaminated-t distribution with shape parameter $\alpha=7$

Contaminated-t distributed errors						
Proc	n	β		θ		Precision
		SE	Conf	SE	Conf	
$\alpha=4$	25	0.00	0.69	0.00	0.87	1.24
	50	0.00	0.65	0.00	0.83	1.25
	100	0.00	0.64	0.00	0.85	1.01
$\alpha=5$	25	0.00	0.67	0.00	0.84	1.70
	50	0.00	0.66	0.00	0.83	1.36
	100	0.00	0.66	0.00	0.84	1.03
$\alpha=6$	25	0.00	0.69	0.00	0.85	1.68
	50	0.00	0.63	0.00	0.83	1.27
	100	0.00	0.64	0.00	0.84	1.64
$\alpha=7$	25	0.00	0.66	0.00	0.85	1.68
	50	0.00	0.63	0.00	0.82	1.42
	100	0.00	0.62	0.00	0.84	1.20
$\alpha=8$	25	0.00	0.66	0.00	0.84	1.54
	50	0.00	0.65	0.00	0.84	1.28
	100	0.00	0.63	0.00	0.82	1.17
$\alpha=9$	25	0.00	0.66	0.00	0.84	1.56
	50	0.00	0.64	0.00	0.81	1.22
	100	0.00	0.66	0.00	0.83	1.15
$\alpha=10$	25	0.00	0.64	0.00	0.83	9.96
	50	0.00	0.58	0.00	0.79	10.98
	100	0.00	0.60	0.00	0.82	6.72
$\alpha=12$	25	0.00	0.61	0.00	0.82	5.40
	50	0.00	0.55	0.00	0.75	4.36
	100	0.00	0.54	0.00	0.77	4.38

The value of precision is computed similarly for $\alpha = 7$ and its neighborhood points. For the shape parameters $\alpha = 10$ and 12 , the precision is significantly high. For highly skewed and contaminated data rank-based method produces a far better fit than REML. Overall precision is high in skewed-t error cases as compared to contaminated error cases. It means empirical efficiencies outperforms REML in a significant magnitude. For $\alpha = 12$, in skew-t distribution precision 28.54 is extremely high than for contaminated error case. It has empirical efficiency 2854% more than REML even with score function $\varphi_7(u)$ that is not appropriate for $\alpha = 12$, the analysis can be further optimized by using $\varphi_{12}(u)$. In the contaminated-t error case, for $\alpha = 10$ the rank method is 1098% more efficient than REML. All the rank-based procedures based on skew-normal score function appeared robust and efficient than REML for skew-t and contaminated-t distribution of error term in multilevel models. But the validity in terms of confidence intervals is not improved in the rank-based method.

5. Summary and Conclusion

The rank-based analysis is a robust and efficient alternative of traditional methods in case of non-normality of error terms in linear models. A suitable selection of score function is necessary to reach optimal analysis. For this purpose, the knowledge about the underlying distribution of the error term is required. Practically, most often Wilcoxon (default) score is used that is suitable for symmetric moderate tailed distribution. For example, rank-based analysis with normal score function applied for normally distributed error terms would produce a completely efficient analysis. In this study, the performance of rank-based analysis based on score function developed for the skew-normal distribution of error in the linear model is evaluated for errors randomly generated from a class of skew-t distributions in multilevel models. And the rank-based approach is compared with the traditional REML in three different sample sizes in terms of validity and efficiency. Two models skew-t and contaminated-t distribution are used with different values of shape parameter α in the family of skew-t distribution. The score function is developed for a particular value of the shape parameter. Thus analysis would be optimized for the correct value of α . We selected the score function appropriate for $\alpha = 7$ and performed rank analysis for $\alpha = 4, 5, 6, 7, 8, 9, 10$ and 12 . The optimized analysis is obtained for β and θ inside 3 elements of correct α in skew-t and contaminated-t distribution. SE and 95% confidence interval for fixed effect parameters are computed. SE is reduced through each rank-based analysis as compared to REML for both fixed effects β and θ , its magnitude also decreases with increasing sample size. Validity in terms of confidence interval should reach a nominal level 0.95, but in a rank-based approach, it is a little bit smaller than REML. Precision is calculated to compare REML and rank-based fit for each case in simulation settings. It is always greater than 1 means rank-based is overall better. For contaminated-t error, precision shows quite similar patterns for $\alpha = 7$ and neighborhood points. But for $\alpha = 10, 12$ with such a high skewness precision is 1098% more efficient than REML. In skewed-t case, the highest value of precision 2854% is attained when. Because REML produces worse fit in such a highly skewed error. Based on the empirical efficiencies, the score function performed well but not for empirical confidence levels.

References

1. Al-shomrani, A. A. (2003). A Comparison of Different Schemes for Selecting and Estimating Score Functions Based on Residuals [Western Michigan University].
2. Arellano-Valle, R. B., & Azzalini, A. (2013). The centred parameterization and related quantities of the skew-t distribution. *Journal of Multivariate Analysis*, 113, 73–90.
3. Auda, H. A., McKean, J. W., Kloke, J. D., & Sadek, M. (2017). A Monte Carlo study of REML and robust rank-based analyses for the random intercept mixed model. *Communications in Statistics: Simulation and Computation*, 48(3), 837–860.
4. Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics*, 12(2), 171–178.
5. Azzalini, A., & Capitanio, A. (2014). *The Skew-Normal and Related Families*. Cambridge University Press.
6. Azzalini, Adelchi. (2013). *The skew-normal and related families*. Cambridge University Press.
7. Azzalini, Adelchi. (2014). *The R sn Package: The skew-normal and skew-t distributions (1.0-0)*.
8. Azzalini, Adelchi, & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2), 367–389.
9. Azzalini, Adelchi, & Genton, M. G. (2008). Robust likelihood methods based on the skew-t and related distributions. *International Statistical Review*, 76(1), 106–129.
10. Basalamah, D. (2017). *Statistical Inference for a New Class of Skew t Distribution and Its Related Properties*. Bowling Green State University.
11. Goldstein, H. (1995). *Multilevel statistical models (second edi)*. Edward Arnold.
12. Hájek, J., & Sidák, Z. (1967). *Theory of rank tests*. Academic Press.
13. Hettmansperger, T. P., & McKean, J. W. (2010). *Robust nonparametric statistical methods (second)*. CRC Press.
14. Jaeckel, L. A. (1972). Estimating Regression Coefficients by Minimizing the Dispersion of the Residuals. *The Annals of Mathematical Statistics*, 43(5), 1449–1458.
15. Kloke, J. D., & McKean, J. W. (2012). Rfit: Rank-based estimation for linear models. *The R Journal*, 4(2), 57–64.
16. Kloke, John D., & McKean, J. W. (2014). *Nonparametric Statistical Methods Using R*. Chapman and

Hall/CRC.

17. Kloeke, John D, McKean, J. W., & Rashid, M. M. (2009). Rank-Based Estimation and Associated Inferences for Linear Models With Cluster Correlated Errors. *Journal of the American Statistical Association*, 104(485), 384–390.
18. McKean, J. W., & Kloeke, J. D. (2014). Efficient and adaptive rank-based fits for linear models with skew-normal errors. *Journal of Statistical Distributions and Applications*, 1(1), 1–18.