Cairo University

The 58th Annual International Conference of Data science

Department of Applied Statistics and Econometrics

8 – 10, Dec, 2025

Index

Performance of double k class estimators in skew-normal mode regression model

Abdelraheim Mohamed Abdelraheim

Department of Basic sciences, Al Madina Higher Institute of Management and Technology, Giza, Egypt,

E-mails: a_mohamedam@yahoo.com

ABSTRACT

The main objective of this article introduces bias, moment matrix, and mean squire error (MSE) with skew-normal mode regression model for the double k class estimators. In previous studies, double k-class estimators were used with symmetric data, whether from a normal or t-distribution. In this study, we use the skew normal distribution, meaning when the data is skewed, and compare the double k-class estimators with the Maximum Likelihood Estimator (MLE). A Monte Carlo simulation study was presented, aiming to compare the double k-class estimators using different values of k1 and k2 with the MLE, using the MSE for comparison. The presented methods are then illustrated through the analysis of real data application.

KEYWORD

Double k class; Skew normal; Mode regression model; Shrinkage estimators

Introduction

Some double k class estimators in the context of multivariate skew-normal errors involve addressing the challenges posed by non-normal error distributions, particularly skewness, in linear regression models. Various other adaptive shrinkage estimators can also be shown as special cases of shrinkage estimators (Ullah et al ,1984).The shrinkage estimator is a generalization of the double k-class estimator, The Stein-rule (SR) estimator and the ordinary least squares (OLS) estimator, often used in econometrics. When errors are skew-normal, traditional estimators may become inefficient or biased. The family of double k-class estimators suggested by Ullah and Ullah (1978) this family was introduced with the assumption of spherical errors and assumption multivariate normal errors. Carter et al (1993) found estimator best performance from the least squares estimator in a general regression model. Vinod and Srivastava (1995) provide of asymptotic bias and mean squared error (MSE) results for the double k-class estimator with nonnormal. Wan and Chaturvedi (2001) introduced the double k-class estimators when the disturbances were non-spherical and the covariance matrix is unknown. The showing of the generalized double k-class estimators with balanced loss function and general pitman closeness were matured by Chaturvedi and Shalabh (2004). The showings of the generalized double k-class estimators with LINEX loss function by Shalabh et al (2012). AbdelRaheim (2016) introduced analysis for the exact and asymptotic properties for double k-class estimator in regression model under multivariate t-errors. Pal et al (2016) introduced shrinkage estimators a special case was taken from him which is the double k class estimators with autoregressive model. Cao et al (2021) present a study that use through it a skew-normal mode regression model and uses a ridge regression estimator.

This article is organized as the following: in section (2) introduce the skew normal mode regrssion model. In section (3) present the main result explain exact and asymptotic for the double k class under skew normal mode regression model. In section (4) conduct Monte Carlo simulations the aim of evaluating performance. In section (5) presents the results of real data applications. The results are discussed in section (6). Prof exact and asymptotic properties are put in Appendix.

2- The skew-normal mode regression model

$y_i \sim SN\left[\mu_i, \sigma^2, \alpha\right]$, where y is a random variable, μ is location parameter, σ^2 is scale parameter, and α is skewness parameter. From Azzalini 1985 the probability density function (pdf) is

$$\underset{sn}{\mathrm{f}}\left(y_i; \mu_i, \sigma^2, \alpha\right) = \frac{2}{\sigma}\varphi\left(\frac{y_i - \mu_i}{\sigma}\right)\varnothing\left(\alpha\left(\frac{y_i - \mu_i}{\sigma}\right)\right) \quad (1)$$

Where $\varphi\left(\frac{y_i - \mu_i}{\sigma}\right)$ is pdf of standard normal and $\varnothing\left(\alpha\left(\frac{y_i - \mu_i}{\sigma}\right)\right)$ is cumulative distribution function (cdf) of the standard normal distribution. The expectation and variance of y $E\left(y_i\right) = \mu_i + \mu_0\left(\alpha\right)\sigma$ and $V\left(y_i\right) = \sigma_0^2\left(\alpha\right)\sigma^2$

Where $\mu_0\left(\alpha\right) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\alpha}{\left(1+\alpha^2\right)^{\frac{1}{2}}}$ and $\sigma_0^2\left(\alpha\right) = 1 - \left(\frac{2}{\pi}\right)\frac{\alpha^2}{\left(1+\alpha^2\right)}$

Azzalini and Capitanio (2014) introduce the mode of the pdf is the following

$$\operatorname{mod} e(y_i) = \mu_i + m_0(\alpha)\sigma \quad (2)$$

Where, $m_0(\alpha) = \mu_0\left(\alpha\right) - \frac{t_0(\alpha)\sigma_0(\alpha) - \operatorname{sgn}(\alpha)e^{\left(-\frac{2\pi}{|\alpha|}\right)}}{2}$ and $t_0(\alpha) = \frac{4-\pi}{2}.\frac{\mu_0^3(\alpha)}{\sigma_0^3(\alpha)}$

The mode regression model when $y_i \sim SN\left[\mu_i, \sigma^2, \alpha\right]$ is $\operatorname{mod} e(y_i) = X_i^T\beta$, see Cao et al (2021).

From Cao et al (2021) used the complete data log likelihood function and the EM algorithm of the MLE the function. The estimated coefficients is

$$\mathrm{b}_{sn} = \left(X^T\hat{W}X\right)^{-1} X^T\hat{W}\mathrm{y} \quad (3)$$

Where, $\hat{W} = diag\left(\exp\left(\hat{\mu}_i\right)\right)$ and $\hat{\mu}_i = X_i^T\hat{\beta} - m_0\left(\hat{\alpha}\right)\hat{\sigma}$

The covariance matrix and the MSE of b_{sn} are the following

$cov\left(\mathrm{b}_{sn}\right) = Var\left(\hat{y}\right)\left(X^T\hat{W}\mathrm{X}\right)^{-1}$, and $MSE\left(\mathrm{b}_{sn}\right) = Var\left(\hat{y}\right)tr\left(X^T\hat{W}\mathrm{X}\right)^{-1}$

Where; $\mathrm{tr}\left(X^T\hat{W}\mathrm{X}\right)^{-1} = \sum_{j=1}^{p}\lambda_j$; λ_j is the jth characteristic root of $\left(X^T\hat{W}\mathrm{X}\right)^{-1}$, $Var\left(\hat{y}\right) = \sigma_0^2\left(\hat{\alpha}\right)\hat{\sigma}^2$ and

$\sigma_0^2\left(\hat{\alpha}\right) = 1 - \frac{2\hat{\alpha}^2}{\pi\left(1+\hat{\alpha}^2\right)}$

The mode regression model is

$$\mathrm{y} = \mathrm{X}\beta + \mathrm{u} \quad (4)$$

Where; $\mathrm{y}_{T\times1}$ is a vector of observations on the dependent variable, $\mathrm{X}_{T\times p}$ is a matrix of observations on explanatory variables, $\beta_{p\times1}$ is a parameter vector, and $\mathrm{u}_{T\times1}$ is a disturbance vector. With the following assumptions: $\mathrm{X}_{T\times p}$ is nonstochastic matrix of rank p, $\mathrm{u}_{T\times1}$ □ $SN\left[\mu = 0, \sigma^2\mathrm{I}, \alpha^*\right]$ and T > P ; μ is location parameter, σ^2 is scale parameter, and $\alpha^* = \alpha l$; α is skewness parameter, l is vector elementally unity

Considering a general class of shrinkage estimators is $b_{sn(sh)} = \left(I + h_1 A\right)^{-1} b_{sn}$. Where; A_{pxp} is any known positive definite matrix and the stochastic scalar h_1 is $h_1 = \frac{k_1 \hat{u}^T \hat{u}}{b^T Bb}$ and $\hat{u} = y - xb$. Where, k_1 is a positive constant and B is a known positive definite matrix. When $A = I$ and $B = X^T X$; $b_{sn(sh)}$ becomes the double k-class estimators is given by

$$\tilde{b}_{sn(k_1,k_2)} = \left[1 - \frac{k_1 u^T u}{y^T y - k_2 u^T u}\right] b_{sn} \tag{5}$$

Where k_1, k_2 are arbitrary scalars which may be stochastic or nonstochastic; $b_{sn(k_1,k_2)}$ represents a family of Double k-class estimators, and James and Stein (1961) introduced the Stein-Rule (SR) estimator $\tilde{b}_{sn(k_1,1)} = \left[1 - \frac{k_1 \hat{u}' \hat{u}}{y' y - \hat{u}' \hat{u}}\right] b_{sn}$

3- THE MAIN RESULTS

The exact and asymptotic distribution of the proposed some estimators has been derived under the assumption that parmeter $\theta = \frac{\left(X\hat{\beta} - m_0(\hat{\alpha})\hat{\sigma}\right)^T \left(X\hat{\beta} - m_0(\hat{\alpha})\hat{\sigma}\right)}{2\,\mathrm{var}(y)}$ the following section illustrates the main results

(3-1) the Exact Results

The sampling error of $\tilde{b}_{sn(k_1,k_2)}$ is $\left(\tilde{b}_{sn(k_1,k_2)} - \beta\right) = \left(b_{sn} - \beta\right) - k_1 c b_{sn}$

Where $c = \frac{y^T M y}{y^T N y}$; $M = I - x\left(X^T \hat{W} X\right)^{-1} X^T \hat{W}$, $N = I - k_2 M$. Further, M is an idempotent matrix with rank n=T-p and N is a nonnegative definite matrix.

Theorem 1: Some exact properties such as the bias vector of estimator, the moment matrix and the MSE of $\tilde{b}_{sn(k_1,k_2)}$ under multivariate skew-normal errors are given by.

Proof of the above theorem is given in the Appendix.

The bias vector of estimator is given by

$$\mathrm{bias}\left(\tilde{b}_{sn(k_1,k_2)}\right) = h - \frac{-nk_1}{2}(\beta + h) g_{2,1} \tag{6}$$

Where; $h = \mu_0(\alpha)\sigma\left(X^T \hat{W} X\right)^{-1} X^T \hat{W} l$; l is vector elementally unity

The moment matrix is given by

$$cov\left(\tilde{b}_{sn(k_1,k_2)}\right) = \mathrm{cov}(b_{sn})\left[1 - nk_1 g_{2,1} - \frac{n(n+2)k_1^2}{4}\left(g_{3,2} - g_{2,2}\right)\right] - \left(\beta\beta^T + 2\beta h^T + hh^T\right)$$
$$\left[nk_1 g_{3,1} + \frac{n(n+2)k_1^2}{4}\left(g_{4,2} - g_{3,2}\right)\right] - nk_1\left(\beta\beta^T + \beta h^T\right) g_{2,1} \tag{7}$$

The MSE is given by

$$MSE\left(\tilde{b}_{sn(k_1,k_2)}\right) = MSE(b_{sn})\left[1 - nk_1 g_{2,1} - \frac{n(n+2)k_1^2}{4}\left(g_{3,2} - g_{2,2}\right)\right] - \left(\beta^T \beta + 2\beta^T h + h^T h\right)$$
$$\left[nk_1 g_{3,1} + \frac{n(n+2)k_1^2}{4}\left(g_{4,2} - g_{3,2}\right)\right] - nk_1\left(\beta^T \beta + \beta^T h\right) g_{2,1} \tag{8}$$

Remark: the exact some properties such as the bias vector of estimator, the moment matrix and the MSE of $\tilde{b}_{sn(k_1,k_2)}$ under multivariate skew-normal error. When let $\alpha = 0$, this leads to the exact some properties such as the bias vector of estimator, the moment matrix and the MSE of $\tilde{b}_{sn(k_1,k_2)}$ under multivariate normal error. Thus, these exact properties in the case $u_{T\times 1} \sim SN\left[\mu, \sigma^2 I, \alpha^*\right]$ are generalized to the same exact properties in the case $u_{T\times 1} \sim N\left[\mu, \sigma^2 I\right]$.

(3-2) the asymptotic expansion under large θ results

The asymptotic distribution of the proposed some estimators has been derived under the assumption large θ with multivariate skew-normal error

Theorem 2: Asymptotic expansion large θ some properties such as the bias vector of estimator, the moment matrix and the MSE of $\tilde{b}_{sn(k_1,k_2)}$ under multivariate skew-normal errors are given by.

Proof of the above theorem is given in the Appendix.

The asymptotic expansion bias vector of estimator up to order $\frac{1}{\theta^3}$ is given by

$$\text{bias}\left(\tilde{b}_{sn(k_1,k_2)}\right) = h - \frac{-nk_1}{2}(\beta + h) g_{2,1}\left[\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left((\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T}\right)\frac{1}{\boldsymbol{\theta}^2}\right] \tag{9}$$

The asymptotic expansion moment matrix up to order $\frac{1}{\theta^3}$ is given by

$$cov\left(\tilde{b}_{sn(k_1,k_2)}\right) = \text{cov}(b_{sn})\left[1 - nk_1\left(\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T}\right]\frac{1}{\boldsymbol{\theta}^2}\right) + \frac{n(n+2)k_1^2}{4\boldsymbol{\theta}^2}\right] - \left(\beta\beta^T + 2\beta h^T + hh^T\right)$$
$$\left[nk_1\left(\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T} - 2\right]\frac{1}{\boldsymbol{\theta}^2}\right) + \frac{n(n+2)k_1^2}{4\boldsymbol{\theta}^2}\right] - nk_1\left(\beta\beta^T + \beta h^T\right)\left(\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T}\right]\frac{1}{\boldsymbol{\theta}^2}\right) \tag{10}$$

The asymptotic expansion MSE up to order $\frac{1}{\theta^3}$ is given by

$$MSE\left(\tilde{b}_{sn(k_1,k_2)}\right) = MSE(b_{sn})\left[1 - nk_1\left(\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T}\right]\frac{1}{\boldsymbol{\theta}^2}\right) + \frac{n(n+2)k_1^2}{4\boldsymbol{\theta}^2}\right] - \left(\beta^T \beta + 2\beta^T h + h^T h\right)$$
$$\left[nk_1\left(\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T} - 2\right]\frac{1}{\boldsymbol{\theta}^2}\right) + \frac{n(n+2)k_1^2}{4\boldsymbol{\theta}^2}\right] - nk_1\left(\beta^T \beta + \beta^T h\right)\left(\frac{1}{\boldsymbol{\theta}} + \frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_2 - \boldsymbol{T}\right]\frac{1}{\boldsymbol{\theta}^2}\right) \tag{11}$$

Remark: also notes that asymptotic expansion large θ some properties such as the bias vector of estimator, the moment matrix, and MSE of $\tilde{b}_{sn(k_1,k_2)}$ under multivariate skew-normal error. When let α = 0, this leads to asymptotic expansion large θ some properties such as the bias vector of estimator, the moment matrix and MSE of $\tilde{b}_{sn(k_1,k_2)}$ under multivariate normal error. Thus, asymptotic expansion large θ these properties in the case $u_{T\times1} \sim SN\left[\mu=0,\sigma^2 I,\alpha^*\right]$ are generalized to the same asymptotic expansion large θ these properties in the case $u_{T\times1} \sim N\left[\mu=0,\sigma^2 I\right]$.

4- SIMULATION STUDY

In this section, we use simulation to evaluate the finite sample performances of the proposed EM algorithm and we compare of some shrinkage estimator such as the double k-class estimators of β with multivariate skew-normal errors under risk function. A simulation study was conducted using the R program. We consider the linear regression model the following compare $y = x\beta + u$. The observations on x are nonstochastic. The error term is generated fond multivariate skew-normal distribution with parameters 0, σ^2 I, α^* where and σ^2 = (4 or 6 or 8), $\alpha^* = \alpha l$, l is vector elementally unity and element skew parameter α=(5 or -5) and (1,2,3,4,5). The number of variables (p), the number of observation is (T): (p, T) = {(4, 50), (4,100)} and the parameter vector β = (1, 2, 3, 4). The data were generated from a multivariate normal distribution, mean 0, Toeplitz correlation matrix and the degrees of multicollinearity ρ=(0.85, 0.90, 0.95, 0.99). the output data are repeated 1000 to evaluate the performance of $\tilde{b}_{sn(k_1,k_2)}$ the estimated of $MSE\left(\tilde{b}_{sn(k_1,k_2)}\right)$ is $\frac{1}{1000}\sum_{i=1}^{1000} MSE\left(\tilde{b}_{sn(k_1,k_2)}\right)$. In the simulations, the following for different values of k_1 and k_2 in the double k-class estimators. Shalabh et al (2012) used the different of double k class estimators the following.

a- The LS estimator; k_1=0.

b- The (SR) estimator; $k_1 = (p-2)/(T-p+2)$ and $k_2 = 1$

c- Minimum mean squared error estimator (MMSE); $k_1 = \left(1+\frac{2}{(T+p)}\right)/\left(T-p+2\right)$ and $k_2 = 1-\frac{1}{T-p}$.

d- Adjusted minimum mean squared error estimator (AMMSE);. $k_1 = \left(1+\frac{2}{(T+p)}\right)/\left(T-p+2\right)$ and $k_2 = 1-\frac{p}{T-p}$

e- Double k-class estimator (KK-C), proposed by Carter et al (1993); $k_1 = (p-2)/(T-p+2)$ and $k_2 = 1-k_1$

These estimators are exposed to criteria, MSE with quadratic loss function under multivariate skew-normal error. The performance of the MLE is assessed by $MSE\left(\tilde{b}_{sn}\right)$ of β under multivariate skew-normal error is defined as $MSE\left(\tilde{b}_{sn(k_1,k_2)}\right)$, The MSE compare double k class estimators with MLE respectively.

The simulation results for some double k-class estimators are presented below.

1- All double k-class estimators are better than the MLE when using the MSE.

2- Using different scales σ, skewness α, and sample sizes, as well as varying degrees of multicollinearity increases, the MSE for the mode of double k-class regression increases.

3- When using different scales σ, skewness α, and degrees of multicollinearity, the performance of the MSE for the mode of double k-class regression is better when the sample size increases.

4- The investigation into the mode of double k-class regression, examining various levels of skewness (α), sample sizes, and multicollinearity, revealed that its MSE performance improves as the scale parameter (σ) decreases see table(2).

5- Across various experimental conditions involving scale (σ), sample size, and multicollinearity, the MSE of the mode of double k-class regression estimator consistently demonstrates superior performance when the skewness parameter α increases. However we note that the MSE improves more effectively at degree of skewness parameter α is 1, 2, 3, but starting from 4, 5, MSE improves less effectively.

6- Generally can say, it is noted that among these estimators, the SR and kk-c estimators are better, whereas the MMSE and AMMSE estimators are worse.

Table 1: The MSE of the double k class estimators when: T=50 and100, σ^2=4

T	α	ρ	MLE	SR	MMSE	AMMSE	KK-C
50	5	0.85	0.09216595	0.05480621	0.07268690	0.07271854	0.05484495
		0.90	0.13512072	0.08904912	0.11109852	0.11113543	0.08909432
		0.95	0.2645793	0.1898137	0.2255944	0.2256511	0.1898831
		0.99	1.2995286	0.9888732	1.1375399	1.1377647	0.9891485
	-5	0.85	0.07468650	0.06563251	0.06996420	0.06996791	0.06563705
		0.90	0.10949514	0.09943082	0.10424584	0.10424984	0.09943572
		0.95	0.2144031	0.2006370	0.2072230	0.2072282	0.2006435
		0.99	1.053408	1.008128	1.029791	1.029808	1.008149
100	5	0.85	0.04483228	0.03286556	0.03872616	0.03873075	0.03287142
		0.90	0.06574047	0.05192131	0.05868902	0.05869404	0.05192774
		0.95	0.1287522	0.1082731	0.1183022	0.1183093	0.1082821
		0.99	0.6341707	0.5568468	0.5947139	0.5947395	0.5568795
	-5	0.85	0.03886666	0.03523065	0.03701106	0.03701173	0.03523151
		0.90	0.05699173	0.05319824	0.05505575	0.05505643	0.05319911
		0.95	0.1116157	0.1070152	0.1092679	0.1092687	0.1070163
		0.99	0.5488147	0.5366574	0.5426103	0.5426124	0.5366601

Table 2: The MSE of the double k class estimators when: T=50, α=5,

σ^2	ρ	MLE	SR	MMSE	AMMSE	KK-C
4	0.85	0.09216595	0.05480621	0.07268690	0.07271854	0.05484495
	0.90	0.13512072	0.08904912	0.11109852	0.11113543	0.08909432
	0.95	0.2645793	0.1898137	0.2255944	0.2256511	0.1898831
	0.99	1.2995286	0.9888732	1.1375399	1.1377647	0.9891485
6	0.85	0.12958551	0.06211661	0.09441725	0.09449735	0.06221471
	0.90	0.1900249	0.1042361	0.1453051	0.1454016	0.1043541
	0.95	0.3721693	0.2276234	0.2968172	0.2969711	0.2278118
	0.99	1.828146	1.205733	1.503669	1.504301	1.206508
8	0.85	0.16874959	0.06182132	0.11302846	0.11319289	0.06202266
	0.90	0.2475129	0.1088510	0.1752506	0.1754526	0.1090984
	0.95	0.4848669	0.2458502	0.3602981	0.3606281	0.2462542
	0.99	2.382015	1.331879	1.834687	1.836072	1.333574

Table 3: The MSE of the double k class estimators when: T=50, σ^2=4,

α	ρ	MLE	SR	MMSE	AMMSE	KK-C
1	0.85	0.16615401	0.03261903	0.09656957	0.09677919	0.03287572
	0.90	0.24368455	0.06847219	0.15237593	0.15263562	0.06879020
	0.95	0.4773295	0.1710654	0.3177154	0.3181443	0.1715905
	0.99	2.3447126	0.9763647	1.6315423	1.6333711	0.9786043
2	0.85	0.12133095	0.06068222	0.08971637	0.08978517	0.06076647
	0.90	0.15214918	0.09786462	0.12384710	0.12389654	0.09792516
	0.95	0.3483969	0.2204526	0.2816966	0.2818262	0.2206113
	0.99	1.711231	1.162891	1.425359	1.425890	1.163541
3	0.85	0.10377547	0.05995056	0.08092763	0.08096984	0.06000225
	0.90	0.15214918	0.09786462	0.12384710	0.12389654	0.09792516
	0.95	0.2979381	0.2093314	0.2517398	0.2518161	0.2094248
	0.99	1.463370	1.092639	1.270071	1.270376	1.093013
4	0.85	0.09609901	0.05728493	0.07586233	0.07589675	0.05732709
	0.90	0.14088939	0.09306545	0.11595424	0.11599436	0.09311457
	0.95	0.2758798	0.1983501	0.2354548	0.2355163	0.1984254
	0.99	1.355028	1.033096	1.187164	1.187408	1.033395
5	0.85	0.09216595	0.05480621	0.07268690	0.07271854	0.05484495
	0.90	0.13512072	0.08904912	0.11109852	0.11113543	0.08909432
	0.95	0.2645793	0.1898137	0.2255944	0.2256511	0.1898831
	0.99	1.2995286	0.9888732	1.1375399	1.1377647	0.9891485

5- Real data application

In this section; we introduce the double k class estimators with skew normal mode regression model by using the C_Reactive_Protein (CRP) data. This dataset includes detailed laboratory test results and diagnostic indicators for 100 patients, collected the dataset from multiple hospitals across the United States within 2021 (https://www.kaggle.com/datasets/taruneshburman/biomarkers-for-cardiovascular-risk-us-hospitals). The CRP data is dependent variable y and four independent variables: x_1 (hdl_cholesterol), x_2 (ldl_cholesterol), x_3 (triglycerides), x_4 (fasting_glucose), x_5 (troponin_level), x_6 (creatinine). We are interested in the relationship between the CRP y and some important variables.

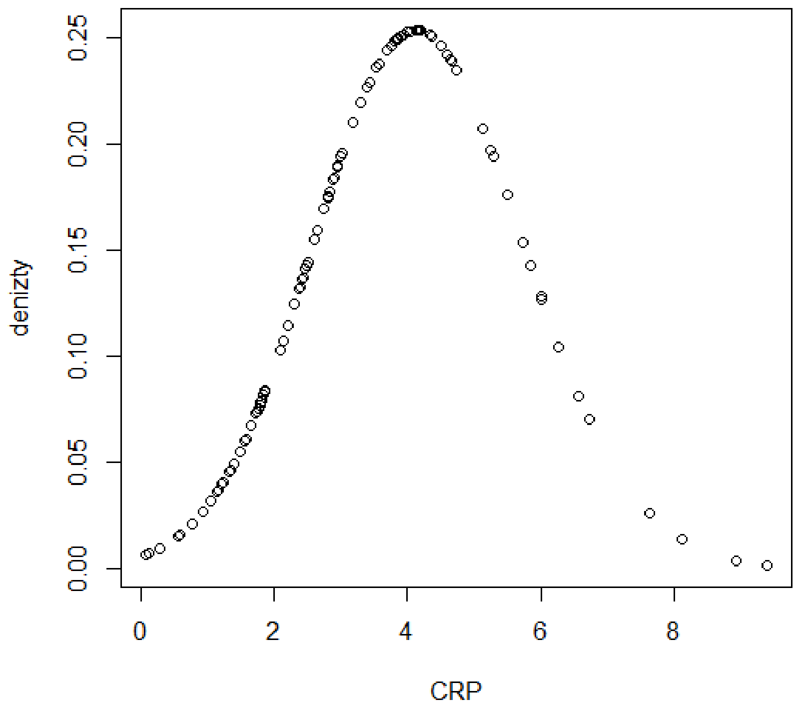

Figure 1: The density estimate of y, C_Reactive_Protein

	X_2	X_3	X_4	X_5	X_6
X_1	0.20	-0.58	-0.12	-0.46	0.28
	X_2	0.89	0.41	0.16	0.33
		X_3	-0.14	-0.93	-0.28
			X_4	-0.37	0.57
				X_5	-0.13

Figure 2: Correlation coefficient of independent variables.

Figure 1 shows that the y distribution skewed, indicating that it approximately follows a skewed normal distribution. As shown in Figure 2, the correlation between the independent variables demonstrates multicolinearity. Based on CRP data, we consider the double k class estimators in skew normal mode regression model.

Table 4: Estimated coefficient, $\hat{\sigma}^2, \hat{\alpha}$ and estimated MSE of CRP

variable	MLE	SR	MMSE	AMMSE	KK-C
X1	0.02132835	0.02225013	0.02156632	0.02158038	0.02229177
X2	-0.00185169	-0.001931717	-0.00187235	-0.001873571	-0.001935332
X3	0.00639112	0.006667337	0.006462431	0.006466644	0.006679813
X4	0.00629774	0.00656992	0.006368008	0.00637216	0.006582215
X5	-2.004841	-2.091488	-2.02721	-2.028532	-2.095402
X6	0.6945715	0.72459	0.7023213	0.7027792	0.725946
$\hat{\sigma}^2$	3.659056	3.817196	3.699882	3.702295	3.824339
$\hat{\alpha}$	0.5710989	0.5710989	0.5710989	0.5710989	0.5710989
MSE	1.1741058	0.3647088	0.9698160	0.9765357	0.3853251

The results are presented Table 4, we observed a different between the double k class and MLE estimation in this dataset, Table 4 shows that the estimation coefficients for the double k class and MLE estimation have the same sign. We find that estimates of variance differ between different estimators and note that the skewness coefficient does not change for different variables. We also, the MSE values for the biased estimators (SR, MMSE, AMMSE, and KK-C) are lower than those of the MLE estimators. The SR and KK-C estimators have the lowest MSE value.

6- CONCLUSION

The main objective of this article is to construct a mode regression model for the skew normal distribution, aiming to identify and study how to apply of double k-class estimators on it. We reviewed four estimators for k1 and k2, then conducted a Monte Carlo simulation study to compare the estimators' results, varying the scales (σ), the skewness (α), the sample sizes T, and degrees of multicollinearity ρ. The Monte Carlo simulation results clearly show that for skewed data, the double k-class estimators consistently outperform the MLE using MSE for comparison across all used values of k1 and k2. Priority should be given to the SR and kk-c estimators as they yield better results than the other estimators. Finally, we recommend using the double k-class estimator with skewed data. This article established a regression modal for the skew normal distribution using double k-class estimators. Among our potential future works is studying double k-class estimators with the skew t-distribution, as well as potentially combination double k-class and ridge estimators.

Appendix

B. proof theorem 1

The G function,

$$g_{i,j} = G_{\left(k_2,\ \theta,\ \frac{T}{2}+i,\ \frac{n}{2}+j\right)} = \int_{-\infty}^{0} h_{\left(t_1, k_2 .\ \theta,\ \frac{T}{2}+i,\ \frac{n}{2}+j\right)} dt_1 \tag{12}$$

Where, $0 \le k_1 \le 1$; i, j = 0, 1, 2, …..

Let $z = \frac{P^T y}{\sqrt{var(y)}} \sim sn\left[\bar{z}, I, \alpha^*\right]$; where P orthogonal matrix, $\bar{z} = \frac{P^T\left(X\hat{\beta} - m_0(\hat{\alpha})\hat{\sigma}\right)}{\sqrt{\mathrm{var}(y)}}$

The first derivative of G function is $\frac{\partial}{\partial \bar{z}} g_{i,j} = \frac{\partial}{\partial \theta} g_{i,j} \frac{\partial \theta}{\partial \bar{z}} = \frac{\partial}{\partial \theta} g_{i,j}\, \bar{z}$

The bias $\tilde{b}_{sn(k_1,k_2)}$ is

$$bias\left(\tilde{b}_{sn(k_1,k_2)}\right) = -k_1 E\left(cb_{sn}\right) \tag{13}$$

Where,

$$E\left(cb_{sn}\right) = \left(X^T \hat{w} X\right)^{-1} X^T \hat{w} \sqrt{var(y)}\ \ P\ E\left(zc\right) \tag{14}$$

And $c = \frac{y^T My}{y^T Ny} = \frac{z^T D_1 z}{z^T D_2 z}$; $P^T MP = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} = D_1$ and $P^T NP = \begin{bmatrix} (1-k_2) I_n & 0 \\ 0 & I_{T-n} \end{bmatrix} = I - k_2 D_1 = D_2$

$$E\left(zc\right) = \bar{z}\frac{\partial}{\partial \theta} E\left(c\right) + \bar{z}E\left(c\right) \tag{15}$$

The bias covariance of $\tilde{b}_{sn(k_1,k_2)}$ is

$$\mathrm{cov}\left(\tilde{b}_{sn(k_1,k_2)}\right) = \mathrm{cov}\left(b_{sn}\right) - 2k_1 E\left(cb_{sn}\, \dot{b}_{sn}^T\right) + 2k_1 E\left(cb_{sn}\right)\beta^T + k_1^2 E\left(c^2 b_{sn}\, \dot{b}_{sn}^T\right) \tag{16}$$

$$E\left(cb_{sn}\, \dot{b}_{sn}^T\right) = Var\left(y\right)\left(X^T \hat{w} X\right)^{-1} X^T \hat{w} P E\left(zz^T c\right) P^T\ \acute{\hat{w}}^T\ X\left(X^T\ \hat{w}X\right)^{-1} \tag{17}$$

$$E\left(zz^T c\right) = \bar{z}\bar{z}^T \frac{\partial^2}{\partial \theta^2} E\left(c\right) + \left(2\bar{z}\bar{z}^T + I\right)\frac{\partial}{\partial \theta} E\left(c\right) + \left(\bar{z}\bar{z}^T + I\right)E\left(c\right) \tag{18}$$

$$E\left(c^2 b_{sn}\, \dot{b}_{sn}^T\right) = Var\left(y\right)\left(X^T \hat{w} X\right)^{-1} X^T \hat{w} P E\left(zz^T c^2\right)\acute{P}\, \hat{w}^T X\left(X^T\ \hat{w}X\right)^{-1} \tag{19}$$

$$E\left(zz^{T}c^{2}\right)=\overline{zz}^{T}\frac{\partial^{2}}{\partial\theta^{2}}E\left(c^{2}\right)+\left(2\overline{zz}^{T}+I\right)\frac{\partial}{\partial\theta}E\left(c^{2}\right)+\left(\overline{zz}^{T}+I\right)E\left(c^{2}\right) \tag{20}$$

The Pdf of multivariate skew-normal of z is $f(z)=2\varphi(z)\varnothing\left(\acute{\alpha}^{*}z\right)$

Where, $\varphi(z)$ is the pdf of multivariate normal of z, $\emptyset(\acute{\alpha}^{*}z)$ is the cdf of multivariate normal of $\acute{\alpha}^{*}z$

The moment generating function of $c=\frac{z^{T}D_{1}z}{z^{T}D_{2}z}$

$$\underset{c}{\mathrm{M}}\left(t_{1},t_{2}\right)=e^{t_{1}z^{T}D_{2}z+t_{2}z^{T}D_{1}z}$$

The first and second derivative is the following

$$\frac{\partial\underset{c}{\mathrm{M}}\left(t_{1},t_{2}\right)}{\partial t_{2}}\Big|_{t_{2}=0}=\frac{n}{2}h_{\left(t_{1},k_{2}.\,\theta,\,\frac{T}{2}+1,\,\frac{n}{2}+1\right)}+\frac{n}{2}\acute{\alpha}^{*}\frac{\partial}{\partial\overline{z}}h_{\left(t_{1},k_{2}.\,\theta,\,\frac{T}{2}+1,\,\frac{n}{2}+1\right)}$$

$$\frac{\partial^{2}\underset{c}{\mathrm{M}}\left(t_{1},t_{2}\right)}{\partial t_{2}^{2}}\Big|_{t_{2}=0}=\frac{n(n+2)}{2}h_{\left(t_{1},k_{2}.\,\theta,\,\frac{T}{2}+2,\,\frac{n}{2}+2\right)}+\frac{n(n+2)}{2}\acute{\alpha}^{*}\frac{\partial}{\partial\overline{z}}h_{\left(t_{1},k_{2}.\,\theta,\,\frac{T}{2}+2,\,\frac{n}{2}+2\right)}$$

The expectation of c and c^2 is the following

$$E(c)=E\left(\frac{z^{T}D_{1}z}{z^{T}D_{2}z}\right)=\int_{-\infty}^{0}\frac{\partial\underset{c}{\mathrm{M}}\left(t_{1},t_{2}\right)}{\partial t_{2}}\Big|_{t_{2}=0}dt_{1} \tag{21}$$

$$E\left(c^{2}\right)=\int_{-\infty}^{0}\frac{\partial^{2}\underset{c}{\mathrm{M}}\left(t_{1},t_{2}\right)}{\partial t_{2}^{2}}\Big|_{t_{2}=0}dt_{1} \tag{22}$$

Substituting from (12), (14), (B.1), (15), (17) into (20) in (13) and(16) we get (6) and (7)

We get (8) follows by applying the trace on both sides of (7)

B. proof theorem 2

large-θ asymptotic expansion

$$\boldsymbol{g}_{2,1}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_{2}-\boldsymbol{T}\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{23}$$

$$\boldsymbol{g}_{1,1}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_{2}-\boldsymbol{T}+2\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{24}$$

$$\boldsymbol{g}_{3,1}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_{2}-\boldsymbol{T}-2\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{25}$$

$$\boldsymbol{g}_{4,1}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+2)\boldsymbol{k}_{2}-\boldsymbol{T}-4\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{26}$$

$$\boldsymbol{g}_{1,2}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+4)\boldsymbol{k}_{2}-\boldsymbol{T}+2\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{27}$$

$$\boldsymbol{g}_{2,2}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+4)\boldsymbol{k}_{2}-\boldsymbol{T}\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{28}$$

$$\boldsymbol{g}_{3,2}=\frac{1}{\boldsymbol{\theta}}+\frac{1}{2}\left[(\boldsymbol{n}+4)\boldsymbol{k}_{2}-\boldsymbol{T}-2\right]\frac{1}{\boldsymbol{\theta}^{2}}+\boldsymbol{O}\left(\frac{1}{\boldsymbol{\theta}^{3}}\right) \tag{29}$$

$$g_{4,2} = \frac{1}{\theta} + \frac{1}{2}\left[(n+4)k_2 - T - 4 \right]\frac{1}{\theta^2} + O\left(\frac{1}{\theta^3}\right) \quad (30)$$

Substituting from (23) into (30) we get (9) and (10). We get (11) follows by applying the trace on both sides of (10).

References

1- AbdelRaheim, A.M., 2016. On Double k-class Estimators With Multivariate t-errors, Acceptance in Pakistan Journal of Statistics, 32, 1-18

2- Azzalini, A. 1985. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics 12 (2), 171-178.

3- Azzalini, A., & Capitanio, A., 2014. The skew normal and related families. New york Cambridge University Press.

4- Cao, X., Wang, D., Wu, L., 2021. Performance of ridge estimator in skew-normal mode regression model, Communications in Statistics – simulation and Computation 52 (3), 1164-1177.

5- Carter, R.A.L., 1985. Double k-class estimators in multiple regression, J. Quant. Econom 1, 27-47.

6- Carter, R.A.L., Srivastava, V.K., Chaturvedi, A., 1993. Selecting a double k-class estimator for regression coefficients, Statistics & Probability Letters 18, 363-371.

7- Chaturvedi, A.,Shalabh., 2004. Risk and Pitman closeness properties of feasible generalized double k-class estimators in linear regression models with non-spherical disturbances under balanced loss function, Journal of Multivariate Analysis 90,229-256.

8- James, W., Stein, C., 1961. Estimation with quadratic loss, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley University of California Press, 361-379.

9- Pal, A. B., Dubey, A. K., Chaturvedi, A., 2016. Shrinkage estimation in spatial autoregressive model. Journal Multivariate Analysis. 143, 362-373.

10- Shalabh., Garg, G., Heumann, C., 2012. Performance of double k-class estimators for coefficients in linear regression models with non-spherical disturbances under asymmetric losses, Journal of Multivariate Analysis 112, 35-47.

11- Ullah, A., Carter, R.A.L., S Srivastava., 1984. The sampling distribution of shrinkage estimators and their F-ratios in the regression model, Journal of Econometrics 25, 109-122.

12- Ullah, A., Ullah, S., 1978. Double k-class estimators of coefcients in linear regression. Econometrica 46, 705–722.

11- Vinod, H. D., & Srivastava, V. K., 1995. Large sample asymptotic properties of the double k-class estimators in linear regression models. Econometric Reviews, 14(1), 75-100.

14- Wan, A.T.K., Chaturvedi, A., 2001. Double k-class estimator in regression models with nonspherical disturbance, Journal of Multivariate Analysis, 79, 226–250.