Available online at www.isas.org.in/jisas

JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 64(2) 2010 237-242

Linear and Nonlinear Approximations of the Ratio of the Standard Normal Density and Distribution Functions for the Estimation of the Skew Normal Shape Parameter

Subir Ghosh* **and Debarshi Dey**

Department of Statistics, University of California, Riverside, CA 92521, USA

Received 02 May 2010; Revised 15 June 2010; Accepted 15 June 2010

SUMMARY

We introduce a linear approximation and a nonlinear approximation of the ratio of the standard normal density and distribution functions in presence of an unknown constant representing the shape parameter of the skew normal distribution. The purpose of these approximations is to estimate the skew normal shape parameter. We present a new estimation method of the shape parameter based on these approximations. The simulation results demonstrate that the approximations strongly resemble their true values in the regions of interest and the estimated biases of the shape parameter are small.

Keywords: Likelihood, Linear approximation, Nonlinear approximation, Skew normal, Standard normal.

1. INTRODUCTION

We consider the ratio R of the standard normal $N(0, 1)$ density function ϕ and the distribution function Φ as

$$R_\lambda(z) = \frac{\phi(\lambda z)}{\Phi(\lambda z)} \tag{1}$$

where λ is a fixed unknown constant. The numerical value of $R_\lambda(0)$ is 0.7978846. Fig. 1 is showing the graphs of $R_\lambda(z)$ against z for $\lambda = \pm 2, \pm 1$, and ± 0.5. The graphs in Fig. 1 intersect at $z = 0$.

For a given λ, we want to approximate $R_\lambda(z)$ by a linear function and a non-linear function for $-3 \le z \le 3$. The approximating functions are

$$A_\lambda(z) = 0.7978846 + \lambda\alpha z \tag{2}$$

$$B_\lambda(z) = 0.7978846 + \lambda\beta(e^{-\delta z} - 1) \tag{3}$$

where α and β are unknown constants and

$$\delta = \begin{cases} +1, & \text{if } \lambda \ge 0 \\ -1, & \text{if } \lambda \le 0 \end{cases}$$

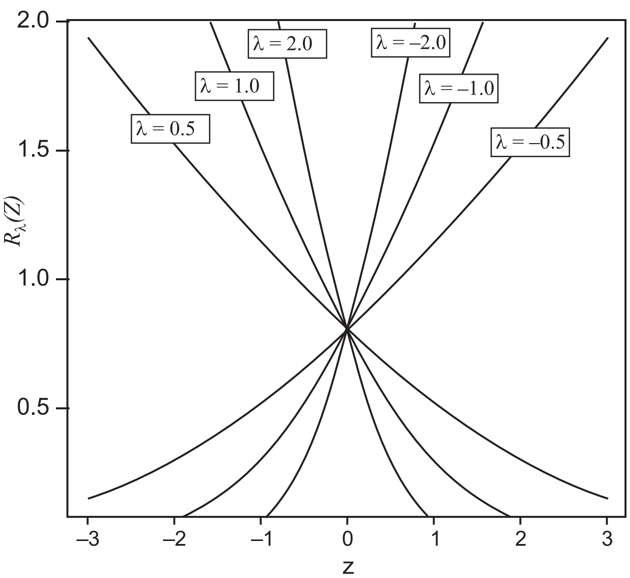

Fig. 1. Plots of $R_\lambda(z)$ against z for $\lambda = \pm 2, \pm 1$, and ± 0.5

Figs. 2 and 3 are counterparts of Fig. 1 for $A_\lambda(z)$ and $B_\lambda(z)$, respectively for the given values of λ and $\lambda\alpha$ in Fig. 2 and the given values of λ and $\lambda\beta$ in Fig. 3. The graphs are now labeled by the values of λ

*Corresponding author : Subir Ghosh
E-mail address : subir.ghosh@ucr.edu

as well as $\lambda\alpha$ in Fig. 2 and $\lambda\beta$ in Fig. 3. Again all the graphs in Figs. 2 and 3 intersect at $z = 0$. The estimation methods of λ, $\lambda\alpha$ and $\lambda\beta$ are given in Section 3.

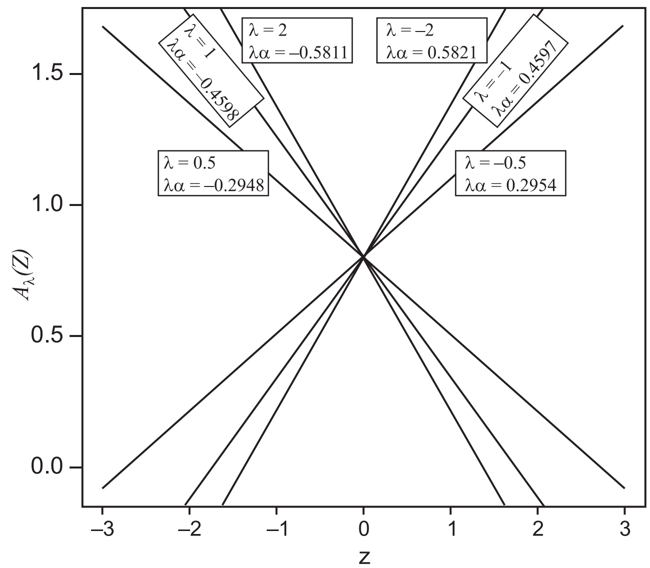

Fig. 2. Plots of $A_\lambda(z)$ against z for $\lambda = \pm2, \pm1$, and ±0.5

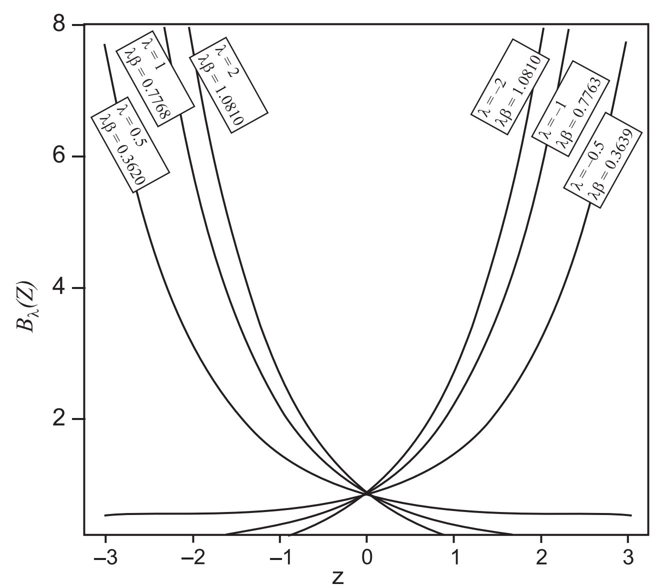

Fig. 3. Plots of $B_\lambda(z)$ against z for $\lambda = \pm2, \pm1$, and $\pm.5$

We want to estimate the unknown parameters λ in (1), $\lambda\alpha$ in (2), and $\lambda\beta$ in (3). In Section 2, we give the motivation for this problem. In Section 3, we present the estimation method for estimating λ, $\lambda\alpha$ and $\lambda\beta$. In Sections 4 and 5, we evaluate the goodness of two approximating functions $A_\lambda(z)$ and $B_\lambda(z)$ for $R_\lambda(z)$ using the simulation method

2. MOTIVATION

The skew normal distribution was introduced in Azzalini (1985, 1986). A random variable Z is said to have a skew normal distribution with the shape parameter λ if its density function at $Z = z$ is

$$f(z;\ \lambda) = 2\phi(z)\Phi(\lambda z) \tag{4}$$

We denote the distribution with the density in (4) as SN(0, 1, λ). When $l = 0$, $\Phi(\lambda z) = \Phi(0) = \frac{1}{2}$, $f(z;\ 0) = f(z)$ and the random variable Z becomes a standard normal variable with the distribution N(0, 1). We now consider a random variable Y for a given value x of a random variable X satisfying

$$Y = \gamma_0 + \gamma_1 x + \sigma Z \tag{5}$$

where γ_0, γ_1, and $\sigma(> 0)$ are unknown fixed constants. The random variable Y is distributed skew normal with density

$$f(y;\ \lambda,\ \gamma_0,\ \gamma_1,\ \sigma) = \frac{2}{\sigma}\phi\left(\frac{y-\gamma_0-\gamma_1 x}{\sigma}\right)\Phi\left(\lambda\frac{y-\gamma_0-\gamma_1 x}{\sigma}\right) \tag{6}$$

We denote the distribution with the density in (6) as SN($\gamma_0 + \gamma_1 x$, σ, λ). We now consider n independent observations (y_i, x_i) from the skew normal distribution with density in (6). We denote $z_i = \frac{y_i - \gamma_0 - \gamma_1 x_i}{\sigma}$, $i = 1, \ldots, n$. The Maximum Likelihood Estimating Equations can be expressed as

$$\sum_{i=1}^{n} z_i = \lambda \sum_{i=1}^{n} R_\lambda(z_i) \tag{7}$$

$$\sum_{i=1}^{n} x_i z_i = \lambda \sum_{i=1}^{n} x_i R_\lambda(z_i) \tag{8}$$

$$\sum_{i=1}^{n} z_i R_\lambda(z_i) = 0 \tag{9}$$

$$\sum_{i=1}^{n} z_i^2 = n \tag{10}$$

We use the equations (7)-(10) to estimate the parameters λ in $R_\lambda(z)$, $\lambda\alpha$ in $A_\lambda(z)$, and $\lambda\beta$ in $B_\lambda(z)$ based on the values of z_i, $i = 1, \ldots, n$.

The complexity of finding the maximum likelihood estimate of λ discussed in details in Sartori (2006) and Dalla Valle (2004). Sartori (2006) provided the modified maximum likelihood estimator with a changed score function to achieve a lower asymptotic bias than the maximum likelihood estimator. We observe that the complexity in the equations (7)-(10) is due to the presence of a complex function $R_\lambda(z)$. We therefore propose a different approach in dealing with the complexity by approximating $R_\lambda(z)$ by $A_\lambda(z)$ in (1) or $B_\lambda(z)$ in (2). It turns out that our approach is much simple to work out.

3. ESTIMATION

We first assume $R_\lambda(z) = A_\lambda(z)$ from (1) and (2). Then from (7)-(9) considering (10) we get

$$\begin{bmatrix} n & \sum_{i=1}^n z_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i z_i \\ \sum_{i=1}^n z_i & n \end{bmatrix} \begin{bmatrix} 0.7978846\,\lambda \\ \lambda^2\alpha \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n z_i \\ \sum_{i=1}^n x_i z_i \\ 0 \end{bmatrix} \qquad (11)$$

We next assume $R_\lambda(z) = B_\lambda(z)$ from (1) and (3). Again from (7)-(9) considering (10) we get

$$\begin{bmatrix} n & \sum_{i=1}^n (e^{-\delta z_i}-1) \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i(e^{-\delta z_i}-1) \\ \sum_{i=1}^n z_i & \sum_{i=1}^n z_i(e^{-\delta z_i}-1) \end{bmatrix} \begin{bmatrix} 0.7978846\,\lambda \\ \lambda^2\beta \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n z_i \\ \sum_{i=1}^n x_i z_i \\ 0 \end{bmatrix} \qquad (12)$$

The equations in (11) and (12) can be expressed in a general form

$$\mathbf{W}\theta = \mathbf{w} \qquad (13)$$

$$\theta = \begin{bmatrix} 0.7978846\lambda \\ \lambda^2\alpha \end{bmatrix} \text{ in (11)}$$

$$\theta = \begin{bmatrix} 0.7978846\lambda \\ \lambda^2\beta \end{bmatrix} \text{ in (12)}$$

$$\mathbf{W} = \begin{bmatrix} n & \sum_{i=1}^n z_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i z_i \\ \sum_{i=1}^n z_i & n \end{bmatrix} \text{ in (11)}$$

$$\mathbf{W} = \begin{bmatrix} n & \sum_{i=1}^n (e^{-\delta z_i}-1) \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i(e^{-\delta z_i}-1) \\ \sum_{i=1}^n z_i & \sum_{i=1}^n z_i(e^{-\delta z_i}-1) \end{bmatrix} \text{ in (12)}$$

$$\mathbf{w} = \begin{bmatrix} \sum_{i=1}^n z_i \\ \sum_{i=1}^n x_i z_i \\ 0 \end{bmatrix} \text{ in both (11) and (12)}$$

Assuming Rank $\mathbf{W} = 2$, the estimates of λ and $\lambda^2\alpha$ from (11) and the estimates of λ and $\lambda^2\beta$ from (12) can be expressed in the general form (Rao (1973))

$$\hat{\theta} = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{w} \qquad (14)$$

From (14), we get the estimates $\hat{\lambda}$ and $\widehat{\lambda\alpha}$ of λ in (1) and $\lambda\alpha$ in (2), respectively when we approximate $R_\lambda(z)$ in (1) by $A_\lambda(z)$ in (2). We also get the estimates $\hat{\lambda}$ and $\widehat{\lambda\beta}$ of λ in (1) and $\lambda\beta$ in (3), respectively when we approximate $R_\lambda(z)$ in (1) by $B_\lambda(z)$ in (3). We denote $A_\lambda(z)$ at $\lambda\alpha = \widehat{\lambda\alpha}$ by $\widehat{A_\lambda}(z)$ and $B_\lambda(z)$ at $\lambda\beta = \widehat{\lambda\beta}$ by $\widehat{B_\lambda}(z)$. For evaluating the goodness of functions in (2) and (3) for approximating $R_\lambda(z)$ in (1), we define

$$\Delta_A(z) = |\, R_\lambda(z) - A_\lambda(z) \,|,\ \Delta_B(z) = |\, R_\lambda(z) - B_\lambda(z) \,| \qquad (15)$$

4. A SIMULATED DATA

We present a simulated data with n = 20, γ_0 = –30, γ_1 = 5, σ = 0.05 and λ = 1. The x_i, i = 1, ..., 20 (Exercise 12.25, Page 522, Mendenhall *et al*. (2009)) are

100, 96, 88, 100, 100, 96, 80, 68, 92, 96, 88, 92, 68, 84, 84, 88, 72, 88, 72, 88

We generate the z_i, i = 1, ..., 20 keeping the nine decimal places. The rounded z_i values at the second decimal place are

1.11, 0.45, –0.26, –0.18, 0.92, 1.48, 1.03, 0.32, 2.03, 1.93, 0.07, 0.67, –0.15, 0.31, –0.11, –1.23, 0.38, 0.01, 1.25, –0.52

We get from (14)

$$\widehat{A_\lambda}(z) = 0.7978846 - 0.3785406z \qquad (16)$$

$$\widehat{B_\lambda}(z) = 0.7978846 + 0.6506317(e^{-z} - 1) \qquad (17)$$

Fig. 4 plots both $R_1(z)$ and $\widehat{A_\lambda}(z)$ against z and Fig. 5 plots both $R_1(z)$ and $\widehat{B_\lambda}(z)$ against z. From Figure 5 and the data considered, we find that the strength of approximation $\widehat{B_\lambda}(z)$ for $R_1(z)$ is very strong within the range of z values between [–1, 1] and moderately strong in the range [–2, –1] and [1, 2]. From Fig. 4 and the data considered, we find that the strength of approximation $\widehat{A_\lambda}(z)$ for $R_1(z)$ is very strong within the range of z values between [–0.5, 2] and moderately strong in the range [–2, –0.5].

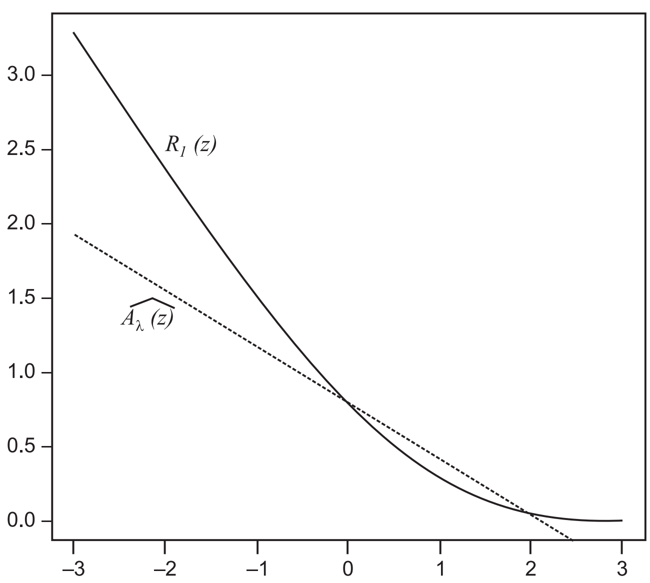

Fig. 4. Plots of both $R_1(z)$ and $\widehat{A_\lambda}(z)$ against z

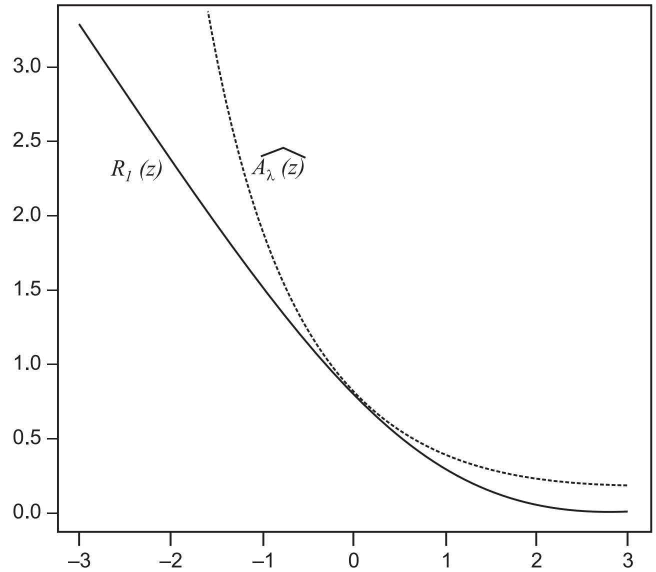

Fig. 5. Plots of both $R_1(z)$ and $\widehat{B_\lambda}(z)$ against z

5. ESTIMATING BIAS AND ACCURACY IN APPROXIMATIONS USING SIMULATIONS

We now obtain 100,000 datasets by repeating the simulation method described in the earlier section 100,000 times. We consider the different values of the parameters γ_0, γ_1, σ and λ. However, we present here the outcomes from only a few values of the parameters. For a set of fixed values of the parameters, we generate 100,000 data sets and from each data set we obtain the numerical values of $\hat{\lambda}$, $\widehat{\lambda^2\alpha}$ or $\widehat{\lambda^2\beta}$ and $\widehat{\lambda\alpha}$ or $\widehat{\lambda\beta}$ for the approximation in (2) or (3) using the equation (14). We calculate also the numerical values of $\Delta_A(z)$ or $\Delta_B(z)$ for the twenty z values in each of 100,000 datasets. We then calculate their average, AveΔ_A or AveΔ_B, over their twenty calculated values. Finally from 100,000 data sets we get 100,000 such AveΔ_A values and 100,000 such AveΔ_B values. We present the First Quartile Q_1, Median, Mean, Third Quartile Q_3, and Standard Deviation (SD) for the 100,000 values of all the cases. In Tables 1- 8, we present only four out of many sets of parameter values that we have done our calculations.

In Tables 1-8, the numerical values of the difference between λ and Median $\hat{\lambda}$ could be considered as the estimates of bias in $\hat{\lambda}$ as an estimate of λ in (1). The numerical values of (λ – Median $\hat{\lambda}$) are –0.0189 in Table 1, –0.0440 in Table 2, –0.0195 in Table 3 and –0.0500 in Table 4 when we estimate λ

Table 1. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\alpha}$, $\widehat{\lambda\alpha}$, and AveΔ_A when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 0.5$, and $\lambda = 0.5$ in approximating (1) by (2)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	0.2858	0.5189	0.6026	0.8255	0.4695
$\widehat{\lambda^2\alpha}$	–0.3190	–0.1567	–0.2291	–0.0570	0.2454
$\widehat{\lambda\alpha}$	–0.3936	–0.2948	–0.2863	–0.1884	0.2938
AveΔ_A	0.0431	0.0857	0.1069	0.1446	0.1815

Table 2. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\alpha}$, $\widehat{\lambda\alpha}$, and AveΔ_A when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 0.5$, and $\lambda = 1.5$ in approximating (1) by (2)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	1.1030	1.5440	1.7120	2.1250	0.8691
$\widehat{\lambda^2\alpha}$	–1.2160	–0.8461	–0.9580	–0.5711	0.5543
$\widehat{\lambda\alpha}$	–0.6101	–0.5540	–0.5473	–0.4767	0.1044
AveΔ_A	0.1078	0.1334	0.1419	0.1660	0.0470

Table 3. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\alpha}$, $\widehat{\lambda\alpha}$, and AveΔ_A when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 1.0$, and $\lambda = 0.5$ in approximating (1) by (2)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	0.2849	0.5195	0.6019	0.8247	0.4682
$\widehat{\lambda^2\alpha}$	–0.3187	–0.1568	–0.2283	–0.0568	0.2441
$\widehat{\lambda\alpha}$	–0.3920	–0.2946	–0.2852	–0.1885	0.1940
AveΔ_A	0.0430	0.0853	0.1050	0.1453	0.1020

Table 4. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\alpha}$, $\widehat{\lambda\alpha}$, and AveΔ_A when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 1.0$, and $\lambda = 1.5$ in approximating (1) by (2)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	1.1080	1.5500	1.7180	2.1340	0.8789
$\widehat{\lambda^2\alpha}$	–1.2180	–0.8477	–0.9609	–0.5728	0.5615
$\widehat{\lambda\alpha}$	–0.6096	–0.5393	–0.5466	–0.4761	0.1044
AveΔ_A	0.1080	0.1336	0.1422	0.1663	0.0473

Table 5. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\beta}$, $\widehat{\lambda\beta}$, and AveΔ_B when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 0.5$, and $\lambda = 0.5$ in approximating (1) by (3)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	0.2585	0.4440	0.5194	0.6828	0.4303
$\widehat{\lambda^2\beta}$	0.0489	0.1598	0.3058	0.3898	0.4355
$\widehat{\lambda\beta}$	0.1793	0.3628	0.3969	0.5905	0.5289
AveΔ_B	0.1056	0.1347	0.1553	0.1834	0.3620

Table 6. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\beta}$, $\widehat{\lambda\beta}$ and AveΔ_B when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 0.5$, and $\lambda = 1.5$ in approximating (1) by (3)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	0.9920	1.4640	1.7520	2.1670	1.1603
$\widehat{\lambda^2\beta}$	0.8752	1.4570	1.8080	2.3260	1.3960
$\widehat{\lambda\beta}$	0.8603	0.9907	0.9647	1.0920	0.1846
AveΔ_B	0.0397	0.0589	0.0737	0.0903	0.0494

Table 7. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\beta}$, $\widehat{\lambda\beta}$ and AveΔ_B when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 1.0$, and $\lambda = 0.5$ in approximating (1) by (3)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	0.2581	0.4444	0.5184	0.6820	0.4274
$\widehat{\lambda^2\beta}$	0.0484	0.1592	0.3040	0.3857	0.4314
$\widehat{\lambda\beta}$	0.1791	0.3618	0.3937	0.5863	0.8121
AveΔ_B	0.1056	0.1347	0.1536	0.1834	0.4250

Table 8. The values of Q_1, Median, Mean, Q_3, and SD for $\hat{\lambda}$, $\widehat{\lambda^2\beta}$, $\widehat{\lambda\beta}$, and AveΔ_B when $\gamma_0 = -30$, $\gamma_1 = 5$, $\sigma = 1.0$, and $\lambda = 1.5$ in approximating (1) by (3)

Estimate	Q_1	Median	Mean	Q_3	SD
$\hat{\lambda}$	0.9934	1.4630	1.7520	2.1700	1.1533
$\widehat{\lambda^2\beta}$	0.8740	1.4500	1.8060	2.3270	1.3910
$\widehat{\lambda\beta}$	0.8583	0.9894	0.9635	1.0910	0.1848
AveΔ_B	0.0396	0.0590	0.0738	0.0908	0.0497

approximating (1) by (2). The numerical values of (λ – Median $\hat{\lambda}$) are 0.0560 in Table 5, 0.0360 in Table 6, 0.0556 in Table 7 and 0.0370 in Table 8 when we estimate λ approximating (1) by (3). The sufficiently small numerical values of the estimated biases thus calculated from Tables 1-8 indicate the absence of an alarming bias in the estimates of λ. We observe that the approximation in (2) provides a small overestimate of λ. On the other hand, the approximation in (3) provides a small underestimate of λ. We also get the similar picture when we use (λ – Mean $\hat{\lambda}$) for estimating the bias in $\hat{\lambda}$ for estimating λ in (1). The median values of AveΔ_A are 0.0857 in Table 1, 0.1334 from Table 2, 0.0853 from Table 3, and 0.1336 from Table 4. The median values of AveΔ_B are 0.1347 in Table 5, 0.0589 in Table 6, 0.1347 in Table 7, and 0.0590 in Table 8. The goodness of two approximating functions in (2) and (3) indicated by the numerical values of AveΔ_A and AveΔ_B are comparable to each other and are sufficiently strong approximating functions of (1).

6. CONCLUSIONS

The complexity for maximum likelihood estimation of the shape parameter λ of the skew normal distribution discussed in Sartori (2006) where the author proposed a modified maximum likelihood estimation. We argue that the source of complexity is from the presence of a complex function $R_\lambda(z)$. The proposed method of estimation of λ in this paper deals with the complexity by approximating $R_\lambda(z)$ by $A_\lambda(z)$ in (1) or $B_\lambda(z)$ in (2). Our simulations demonstrate that the approximations (2) and (3) perform satisfactorily in the regions of interest and the biases in the estimates of λ using the approximations (2) and (3) are in fact sufficiently small.

REFERENCES

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statist*., **12**, 171-178.

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

Dalla Valle, A. (2004). The skew-normal distribution. In: *Skew-Elliptical Distributions and their Applications*, Ed. Genton, M.G., 3-24. Chapman & Hall/CRC: Boca Raton, Florida.

Mendenhall, W., Beaver, R.J. and Beaver, B.M. (2009). *Introduction to Probability* and *Statistics*. 13th edition. Cengage/Duxbury, Florence, Kentucky.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications*. 2nd edition. Wiley, New York.

Sartori, N. (2006). Bias prevention of maximum likelihood scalar skew normal and skew t distributions. *J. Statist. Plann. Inf.,* **136**, 4259-4275.