International Journal of Advanced Statistics and Probability, 3 (1) (2015) 83-92
www.sciencepubco.com/index.php/IJASP
©Science Publishing Corporation
doi: 10.14419/ijasp.v3i1.4610
Research Paper

On the skewed sinh-normal distribution

S.K. Ashour, A.M. Saad*

Department of Mathematical Statistics, Institute of Statistical Studies & Research, Cairo University, Egypt
**Corresponding author E-mail: Ahmusaad@gmail.com*

Copyright © 2015 A.M. Saad, S.K. Ashour. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Leiva et al. [10] introduce the skewed sinh-normal distribution, which is a skewed version of the sinh-normal distribution, discussed some of its properties and characterized an extension of the Birnbaum–Saunders distribution associated with this distribution. In this paper, we will introduce further properties of the skewed sinh-normal distribution, and introduce a new approximate form of its probability density function and cumulative distribution function (cdf), along with numerical comparison between the exact and approximate values of its cdf. Moreover, a random number generator for the distribution will be suggested and a goodness of fit test will be carried out to examine the effectiveness of the proposed random number generator. Finally, maximum likelihood estimation for the unknown parameter of the skewed sinh-normal distribution will be investigated, and numerical illustration for the new results will be discussed.

Keywords: *Skew Distributions; Approximation; Random Number Generator; Maximum Likelihood Estimation.*

1. Introduction

Although the normal distribution is still the most commonly used distribution in both statistical theory and application, it has been observed in various applications that data do not conform to the normal distribution or any other symmetric distribution, which built strong motivation to obtain more flexible probability models that admit asymmetry. In his attempts to build non-normal distributions, Johnson [7] mentioned that it is natural and also convenient to build non-normal distributions by transforming a random variable (RV) that follows a normal distribution. He used the translation method to generate models of probability which could assume wide varieties of shapes by the transformation

$$Z = \nu + \delta g(Y; \gamma, \sigma) \tag{1}$$

Where the variate $Z \sim N(0,1)$ and $g(Y; \gamma, \sigma)$ is a monotone function. Later, Rieck [12] defined that, if $\delta = 2/\alpha$ and $g(Y; \gamma, \sigma) = \sinh(y - \gamma/\sigma)$ in (1) then Y follow a four parameters sinh-normal (SHN) distribution with shape parameters $\nu \in \mathbb{R}$, $\alpha > 0$, location parameter $\gamma \in \mathbb{R}$, and scale parameter $\sigma > 0$, and the notation $Y \sim SHN(\alpha, \gamma, \sigma, \nu)$ is used, which is reduced simply to $Y \sim SHN(\alpha, \gamma, \sigma)$ when $\nu = 0$; for more details and applications of the SHN distribution see Rieck and Nedelman [13], Galea et al. [4], and Leiva et al. [9]. If $Y \sim SHN(\alpha, \gamma, 2)$, then $T = \exp(Y)$ follows the Birnbaum-Saunders (BS) distribution with shape parameter $\alpha > 0$ and scale parameter $\beta = \exp(\gamma) > 0$, which is denoted by $T \sim BS(\alpha, \beta)$; see Birnbaum and Saunders [3], Johnson et al. [8], and Sanhueza et al. [14]. For this reason, the SHN distribution is also called the log-Birnbaum-Saunders distribution.

In analogous way to the definition of the SHN distribution, Leiva et al. [10] introduce a skewed version of the SHN distribution, which they call the skewed sinh-normal (SSN) distribution, by replacing the standard normal distribution by the skew-normal (SN) distribution introduced by Azzalini [2], i.e. they consider the transformation:

$Z = \nu + \frac{2}{\alpha}\sinh\left(\frac{Y-\gamma}{\sigma}\right)$, where $Z \sim SN(\lambda)$ and $\lambda \in \mathbb{R}$ is the shape parameter which determine the skewness.

Now the notation $Y\sim SNN(\alpha,\gamma,\sigma,\nu,\lambda)$ is used, where $\alpha>0$, $\nu\in\mathbb{R}$ and $\lambda\in\mathbb{R}$ are shape parameters, $\gamma\in\mathbb{R}$ is the location parameter, and $\sigma>0$ is the scale parameter. The probability density function (pdf) of SNN distribution is given by:

$$f_Y(y)=\frac{4}{\alpha\sigma}\cosh\left(\frac{y-\gamma}{\sigma}\right)\phi\left(\nu+\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right)\Phi\left(\lambda\left[\nu+\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right]\right),y\in\mathbb{R}$$

And, its cumulative distribution function (cdf) is given by:

$$F_Y(y)=\Phi_\lambda\left(\nu+\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right),y\in\mathbb{R}$$

Where $\Phi_\lambda(\cdot)$ is the cdf of the SN distribution, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively. In addition to introducing the SSN distribution, Leiva et al. [10] characterize its corresponding extension of the Birnbaum distribution as follows: if $Y\sim SNN(\alpha,\gamma,\sigma,\nu,\lambda)$ then the rv $T=\exp(Y)$ follow extended Birnbaum-Saunders distribution with shape parameters $\alpha>0$, $\nu\in\mathbb{R}$ and $\lambda\in\mathbb{R}$, scale parameter $\beta=\exp(\gamma)>0$ and power parameter $\sigma>0$, its pdf is given by

$$f_T(t)=\frac{2\left[t^{2/\sigma}+\beta^{2/\sigma}\right]}{\left[\sigma\alpha\beta^{1/\sigma}t^{\left(\frac{1}{\sigma}+1\right)}\right]}\phi\left(\nu+\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{1/\sigma}-\left(\frac{\beta}{t}\right)^{1/\sigma}\right]\right)\Phi\left(\nu+\frac{\lambda}{\alpha}\left[\left(\frac{t}{\beta}\right)^{1/\sigma}-\left(\frac{\beta}{t}\right)^{1/\sigma}\right]\right);t>0$$

Where, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, and the notation $T\sim EBS(\alpha,\beta,\sigma,\nu,\lambda)$ is used.
The main purposes of this article are: (1) introduce further properties of the SSN distribution, (2) introduce a new approximate form of the pdf and cdf of the SNN distribution, and study the accuracy of the proposed approximate form through a numerical comparison between the exact and the approximate values to its cdf. (3) introduce a random number generator for the SNN distribution, (4) maximum likelihood (ML) estimation for the unknown parameter of the SSN distribution, and (5) numerical illustration for the new results which contains, first, a goodness of fit test for random samples generated using the proposed random number generator in order to assess its effectiveness, second, study the performance of the ML estimates of the SSN distribution depending on random sample generated using the proposed random number generator, for different sample sizes and different values of the unknown parameter.

2. Further properties of the SSN distribution

In this section, we will introduce several theorems about the properties of the skewed sinh-normal distribution.

Theorem 1: *If $Y\sim SNN(\alpha,\gamma,\sigma,\nu,\lambda)$ then its cumulative distribution function is given by*

$$F_Y(y;\alpha,\gamma,\sigma,\nu,\lambda)=\Phi_\lambda(\nu+b_y)$$

Where $b_y=2\sinh([y-\gamma]/\sigma)/\alpha$, $\Phi_\lambda(\cdot)$ is the cdf of $SN(\lambda)$, and it satisfies the following properties:
Property (1): $F_Y(y;\alpha,\gamma,\sigma,\nu,1)=\left(\Phi(\nu+b_y)\right)^2$,
Property (2): $F_Y(y;\alpha,\gamma,\sigma,\nu,\lambda)=1-F_Y(-y;\alpha,-\gamma,\sigma,-\nu,-\lambda)$,
Property (3): $F_Y(y;\alpha,\gamma,\sigma,\nu,\lambda)=\Phi(\nu+b_y)-2T(\nu+b_y,\lambda)$.
Where $\Phi(\cdot)$ is the cdf of the standard normal distribution, and $T(x,h)$ is the function studied by Owen [11].

Theorem 2: *If $\sim SNN(\alpha,\gamma,\sigma,\nu,\lambda)$, where $\alpha>0$, $\nu\in\mathbb{R}$, $\lambda\in\mathbb{R}$, $\gamma\in\mathbb{R}$ and $\sigma>0$ then the general form for the rth population moment about zero is given by*

$$E(Y^r)=\sum_{k=0}^{r}\binom{r}{k}w_k\gamma^{r-k},\ w_k=\int_{-\infty}^{\infty}2\sigma^k\left\{\sinh^{-1}\left[\frac{\alpha}{2}(t-\nu)\right]\right\}^k\phi(t)\Phi(\lambda t)dt$$

Where, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively.

Theorem 3: *If $Y\sim SNN(\alpha,\gamma,\sigma,\nu,\lambda)$, then $-Y\sim SNN(\alpha,-\gamma,\sigma,-\nu,-\lambda)$.*

Theorem 4: *If $Y\sim SNN(\alpha,\gamma,\sigma,\nu,\lambda)$, then any linear combination of Y still have skewed sinh-normal distribution. Moreover if $V=a+bY$ where $a\in\mathbb{R}$, then for $b>0$ $V\sim SNN(\alpha,a+b\gamma,b\sigma,\nu,\lambda)$, and $V\sim SNN(\alpha,a+b\gamma,|b|\sigma,-\nu,-\lambda)$ for $b<0$.*

Theorem 5: *If* $Y \sim SNN(\alpha, \gamma, \sigma, \nu, \lambda)$, *then* $Y_\circ = 2(Y-\gamma)/\alpha\sigma$ *converges to skew normal distribution with parameter* λ *when* $\alpha \to 0$.

Note: the proof of theorems 1 and 2 are easily obtained using simple calculations on the pdf of SSN distribution, while the proof of theorems 2, 3 and 4 can be readily seen by change of variable theorem.

3. A new approximation form for the SSN distribution

Although that the SN and SSN distributions share the flexibility of modeling skewed data they also have an important disadvantage in common, since the SN and SSN distribution pdf contain the cdf of the standard normal distribution, and hence they are not in a closed form. Using Hoyt [6] approximation, Ashour and Abdel-hamed [1] introduce an approximate formula $h(x,\lambda)$ for the skew normal pdf, this approximation is easier for calculations, convenient and in closed form for any value of the rv X, and the skew parameter $\lambda \neq 0$, they also drive the corresponding approximation of the skew normal cdf $H(x,\lambda)$, they also make a comparison between the approximated cdf $H(x,\lambda)$ and the exact cdf of the skew normal distribution to ensure the accuracy of the approximation, and they show that the maximum absolute error is not exceeding 7.0604×10^{-3} when λ varies from 0 to 5 , and when λ varies from 5 to 10 the maximum absolute error increase from 7.0604×10^{-3} to 0.0519.
Now, we will define an approximation to the SSN distribution pdf and cdf using Ashour and Abdel-hamed [1] approximation by establishing the relationship between the SSN distribution and the SN distribution which follow that: if a rv Z have SN distribution with parameter λ, then the rv Y defined as

$$Y = \gamma + \sigma\, sinh^{-1}\left(\frac{\alpha[Z-\nu]}{2}\right)$$

Follow SSN distribution with parameters $\alpha, \gamma, \sigma, \nu,$ and λ, for simplicity, and without loss of generality, we will set $\gamma = 0$ and $\sigma = 1$, hence the pdf of Y is

$$f_Y(y) = \frac{4}{\alpha} cosh(y)\, \phi\left(\nu + \frac{2}{\alpha} sinh(y)\right) \Phi\left(\lambda\left[\nu + \frac{2}{\alpha} sinh(y)\right]\right);\ y \in \mathbb{R}$$

$$= 2b'_y \phi(\nu + b_y)\Phi(\lambda[\nu + b_y]).$$

Where, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and is cdf of the standard normal distribution and $b_y = 2\, sinh(y)/\alpha$, $b'_y = 2\, cosh(y)/\alpha$ is the first derivative of b_y.
Hence, the pdf of the SSN distribution could be written in terms of the SN distribution pdf as follows:

$$f_Y(y) = b'_y \phi_\lambda(\nu + b_y);\ y \in \mathbb{R} \tag{2}$$

Where $\phi_\lambda(\cdot)$ is the pdf of the SN distribution, i.e., $\phi_\lambda(z) = 2\, \phi(z)\Phi(\lambda z); z \in \mathbb{R}$.
Now, our proposed approximation for the SSN distribution is yield by replacing the pdf of the SN distribution in (2), i.e., $\phi_\lambda(\nu + b_y)$ by its approximation i.e., $h(\nu + b_y, \lambda)$ we will denote it by $h(y,\theta)$, where $\theta = (\alpha, \nu, \lambda)$, hence

$$h(y,\theta) = b'_y h(\nu + b_y, \lambda);\ y \in \mathbb{R}$$

Therefore, the approximate pdf of the SNN distribution is given by

$$h(y,\theta) = \begin{cases} 0\,; y < sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda + 3)\right) \\ h_1(y,\theta)\,; sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda + 3)\right) \le y < sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda + 1)\right) \\ h_2(y,\theta)\ ; sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda + 1)\right) \le y < sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda - 1)\right) \\ h_3(y,\theta)\,; sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda - 1)\right) \le y < sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda - 3)\right) \\ h_4(y,\theta)\,; y \ge sinh^{-1}\left(-\frac{\alpha}{2\lambda}(\nu\lambda - 3)\right) \end{cases}$$

Where

$$h_1(y,\theta)=\frac{b'_y}{8\sqrt{2\pi}}e^{-(v+b_y)^2/2}\left(9\lambda(v+b_y)+3\lambda^2(v+b_y)^2+\frac{\lambda^3(v+b_y)^3}{3}+9\right),$$

$$h_2(y,\theta)=\frac{b'_y}{4\sqrt{2\pi}}e^{-(v+b_y)^2/2}\left(3\lambda(v+b_y)-\frac{\lambda^3(v+b_y)^3}{3}+9\right),$$

$$h_3(y,\theta)=\frac{b'_y}{8\sqrt{2\pi}}e^{-(v+b_y)^2/2}\left(9\lambda(v+b_y)-3\lambda^2(v+b_y)^2+\frac{\lambda^3(v+b_y)^3}{3}+7\right),$$

And

$$h_4(y,\theta)=b'_y\sqrt{\frac{2}{\pi}}e^{-(v+b_y)^2/2}.$$

Now, the corresponding approximate formula for the SSN cdf, $H(y,\theta)$ is integrate out $h(y,\theta)$ up to y, which need a lot of simple but rather tedious calculations, and it is given as follows

$$H(y,\theta)=\begin{cases}0\,;y<sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda+3)\right)\\ H_1(y,\theta)\,;sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda+3)\right)\le y<sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda+1)\right)\\ H_2(y,\theta)\;;sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda+1)\right)\le y<sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda-1)\right)\\ H_3(y,\theta)\,;sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda-1)\right)\le y<sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda-3)\right)\\ 2\Phi(v+b_y)-1\,;y\ge sinh^{-1}\left(-\frac{\alpha}{2\lambda}(v\lambda-3)\right)\end{cases}$$

Where

$$H_1(y,\theta)=\frac{1}{8\sqrt{2\pi}}\Bigg[(3\lambda^2+9)\int_{-3/\lambda}^{v+b_y}e^{-\frac{y^2}{2}}dy+\lambda\left(3+\frac{2}{3}\lambda^2\right)e^{-\frac{9}{\lambda^2}}$$

$$-\lambda\left(9+\frac{2}{3}\lambda^2+3\lambda(v+b_y)+\frac{1}{3}\lambda^2(v+b_y)^2\right)e^{-\frac{(v+b_y)^2}{2}}\Bigg],$$

$$H_2(y,\theta)=\frac{1}{8\sqrt{2\pi}}\Bigg[(3\lambda^2+9)\int_{-3/\lambda}^{-1/\lambda}e^{-\frac{y^2}{2}}dy+8\int_{-1/\lambda}^{v+b_y}e^{-\frac{y^2}{2}}dy+\lambda\left(3+\frac{2}{3}\lambda^2\right)e^{-\frac{9}{\lambda^2}}$$

$$-\lambda(1+2\lambda^2)e^{-\frac{1}{\lambda^2}}+\frac{2}{3}\lambda\left(-9+2\lambda^2+\lambda^2(v+b_y)^2\right)e^{-\frac{(v+b_y)^2}{2}}\Bigg],$$

And

$$H_3(y,\theta)=\frac{1}{8\sqrt{2\pi}}\Bigg[(3\lambda^2+9)\int_{-3/\lambda}^{-1/\lambda}e^{-\frac{x^2}{2}}dx+8\int_{-1/\lambda}^{1/\lambda}e^{-\frac{x^2}{2}}dx+(7-3\lambda^2)\int_{1/\lambda}^{v+b_y}e^{-\frac{y^2}{2}}dy$$

$$+\lambda\left(3+\frac{2}{3}\lambda^2\right)e^{-\frac{9}{\lambda^2}}-\frac{1}{3}\lambda\left(27-9\lambda x+2\lambda^2+\lambda^2(v+b_y)^2\right)e^{-\frac{(v+b_y)^2}{2}}\Bigg].$$

A computer program using Mathcad package has been used to present tables of the approximated cdf of the SSN distribution, for y ranging from 0.00 to 2.00 with increment 0.01 for different values of the parameters α, v and λ, while the location and scale parameters γ and σ are fixed to be 0 , and 2 respectively (for example see table 1). Moreover, to study the accuracy of the proposed approximation of the SSN distribution, another computer program using Mathcad package has been used to obtain the absolute difference between the value of approximated and exact cdf in the range of y, and all the results is listed in table 2.

Table 1: The Approximate Cdf of the SSN Distribution for Some Values of the Parameters α, ν and λ.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
(a)	$\alpha = 0.5, \nu = 0, \lambda = 1$									
0.0	0.26	0.27	0.29	0.31	0.32	0.34	0.36	0.38	0.40	0.42
0.1	0.43	0.45	0.47	0.49	0.51	0.53	0.55	0.57	0.59	0.61
0.2	0.63	0.64	0.66	0.68	0.70	0.71	0.73	0.75	0.76	0.78
0.3	0.79	0.80	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89
0.4	0.90	0.91	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96
0.5	0.96	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.99	0.99
0.6	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00
0.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(b)	$\alpha = 1.5, \nu = -1, \lambda = 1$									
0.0	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.04
0.1	0.04	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05
0.2	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.07	0.08
0.3	0.08	0.08	0.09	0.09	0.09	0.09	0.10	0.10	0.10	0.11
0.4	0.11	0.11	0.12	0.12	0.12	0.13	0.13	0.14	0.14	0.15
0.5	0.15	0.15	0.16	0.16	0.17	0.17	0.18	0.18	0.19	0.19
0.6	0.20	0.20	0.21	0.22	0.22	0.23	0.23	0.24	0.25	0.25
0.7	0.26	0.27	0.27	0.28	0.29	0.30	0.30	0.31	0.32	0.33
0.8	0.33	0.34	0.35	0.36	0.37	0.38	0.38	0.39	0.40	0.41
0.9	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.48	0.49	0.50
1.0	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59	0.60
1.1	0.61	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.69	0.70
1.2	0.71	0.72	0.73	0.74	0.75	0.76	0.77	0.78	0.79	0.80
1.3	0.80	0.81	0.82	0.83	0.84	0.84	0.85	0.86	0.87	0.87
1.4	0.88	0.89	0.89	0.90	0.90	0.91	0.92	0.92	0.93	0.93
1.5	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.97
1.6	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.99	0.99
1.7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00
1.8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(c)	$\alpha = 1, \nu = 1, \lambda = 0.5$									
0.0	0.76	0.76	0.77	0.78	0.78	0.79	0.80	0.80	0.81	0.81
0.1	0.82	0.82	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.87
0.2	0.87	0.87	0.88	0.88	0.89	0.89	0.90	0.90	0.90	0.91
0.3	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.93	0.94	0.94
0.4	0.94	0.94	0.95	0.95	0.95	0.95	0.96	0.96	0.96	0.96
0.5	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.98	0.98	0.98
0.6	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.99
0.7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
0.8	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
(d) $\alpha = 1, \nu = 1, \lambda = 2$										
0.0	0.68	0.69	0.70	0.71	0.72	0.73	0.74	0.75	0.75	0.76
0.1	0.77	0.78	0.79	0.79	0.80	0.81	0.81	0.82	0.83	0.83
0.2	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.88	0.88	0.89
0.3	0.89	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.92	0.93
0.4	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.95	0.96
0.5	0.96	0.96	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.98
0.6	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99
0.7	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
0.8	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2: The Maximum Absolute Error of the Proposed Approximation for Some Values of λ.

Value of λ	Maximum Absolute Error
0.1	1.871×10^{-3}
0.2	3.524×10^{-3}
0.4	5.684×10^{-3}
0.8	6.376×10^{-3}
1	5.879×10^{-3}
2	3.267×10^{-3}
5	1.206×10^{-3}
10	0.044

When we study the maximum absolute error for different values of the SSN parameters α, ν and λ, we find that the shape parameters α and ν, does not potentially affect the accuracy of the approximation ,so we fix their values. We conclude that the maximum absolute error increase with λ, when $0 < \lambda \leq 0.8$ to reach 6.376×10^{-3} at $\lambda = 0.8$, and then decreases as λ decrease to reach its lowest value at $\lambda = 6$, and then increase again, but in general the maximum absolute error does not exceeding 0.044, when $6 < \lambda \leq 10$.

4. Random number generator for the SSN distribution

As we mentioned in section (2) a rv Y follows SSN distribution with shape parameters $\alpha > 0, \nu \in \mathbb{R}$ and $\lambda \in \mathbb{R}$, location parameter $\gamma \in \mathbb{R}$, and scale parameter $\sigma > 0$, is defined by

$$Y = \gamma + \sigma \, sinh^{-1}\left(\frac{\alpha[Z-\nu]}{2}\right) \tag{3}$$

Where $Z \sim SN(\lambda)$. This relationship could be used to generate random numbers from $Y \sim SNN(\alpha, \gamma, \sigma, \nu, \lambda)$, since if we have an appropriate random number generator for $Z \sim SN(\lambda)$, for desired values of α, γ, σ, and ν, therefore (3) defines a random number generator for the SSN distribution. several representation of the SN random variates in terms of normal distribution, truncated normal each of which could be used to generate SN random numbers, but we shall use the representation introduced by Henze [5], which state that if U_1 and U_2 are independent uniformly (0,1) distributed random variables then

$\delta\sqrt{-2\,ln\,U_1}\;\; sin(\pi U_2) + \sqrt{1-\delta^2}\,\sqrt{-2\,ln\,U_1}\;\; cos(\pi U_2) \sim SN(\lambda)$, Where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$

Therefore, the above representation of the SN distribution will be used, since it depends on random numbers from the uniform distribution, which is too easy to be generated using computer programs. Thus the algorithm steps to generate random numbers form SSN distribution will be as follows:

1) Generated random numbers u_1 and u_2 from two independent uniformly (0, 1) distributed random variable.
2) Set value for λ, then $z = \delta\sqrt{-2\ln u_1}\ \sin(\pi u_2) + \sqrt{1-\delta^2}\ \sqrt{-2\ln u_1}\ \cos(\pi u_2)$ is a random number form $Z \sim SN(\lambda)$, where $\delta = \lambda/\sqrt{1+\lambda^2}$.
3) Set values for α, γ, σ, and ν, hence $Y = \gamma + \sigma \sinh^{-1}(\alpha[Z-\nu]/2)$ is a random number from $Y \sim SNN(\alpha, \gamma, \sigma, \nu, \lambda)$.

Obviously, the above algorithm is quite simple and easy to be performed using a computer program; however, a goodness of fit test is needed in order to assess its effectiveness, which will be discussed in section (6).

5. Maximum likelihood estimation for the unknown parameters of the SSN distribution

Next, we will discuss ML estimation of the parameters $\alpha, \gamma, \sigma, \nu$ and λ of the SSN distribution. Suppose that $Y_1, Y_2, \dots, Y_n$ are a random sample of size n from the SSN distribution with α, ν and λ shape parameters and γ location parameters, σ is a scale parameter, then the pdf of each Y_i is given by

$$f_{Y_i}(y_i) = \frac{4}{\alpha\sigma} cosh\left(\frac{y_i-\gamma}{\sigma}\right) \phi\left(\nu + \frac{2}{\alpha} sinh\left(\frac{y_i-\gamma}{\sigma}\right)\right) \Phi\left(\lambda\left[\nu + \frac{2}{\alpha} sinh\left(\frac{y_i-\gamma}{\sigma}\right)\right]\right)$$

To simplify the calculations on the pdf we will take $b_{y_i} = 2\, sinh([y_i-\gamma]/\sigma)/\alpha$, and then the first derivative of b_{y_i} is $b'_{y_i} = 2\, cosh([y_i-\gamma]/\sigma)/\alpha\,\sigma$, hence

$$f_{Y_i}(y_i) = 2b'_{y_i}\phi(\nu + b_{y_i})\Phi(\lambda[\nu + b_{y_i}]);\ y \in \mathbb{R}, i = 1,2,\dots,n$$

Then the likelihood function is given by

$$L(\alpha, \gamma, \sigma, \nu, \lambda) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \prod_{i=1}^{n} 2b'_{y_i}\phi(\nu + b_{y_i})\Phi(\lambda[\nu + b_{y_i}])$$

And, the log likelihood function will be

$$ln\, L(\alpha, \gamma, \sigma, \nu, \lambda) \propto \sum_{i=1}^{n}\{ln\, \phi(\nu + b_{y_i}) + ln\, \Phi(\lambda[\nu + b_{y_i}]) + ln\, b'_{y_i}\} \quad (4)$$

Differentiating (4) with respect to the parameters $\alpha, \gamma, \sigma, \nu$ and λ, and then equating each derivative with zero will gives the likelihood equations as follow

$$\sum_{i=1}^{n}\left\{\frac{1}{\hat{\alpha}}(\hat{\nu} + \widehat{b_{y_i}})\widehat{b_{y_i}} - \frac{\hat{\lambda}}{\hat{\alpha}}\frac{\phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}{\Phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}\widehat{b_{y_i}} - \frac{1}{\hat{\alpha}}\right\} = 0,$$

$$\sum_{i=1}^{n}\left\{(\hat{\nu} + \widehat{b_{y_i}})\widehat{b'_{y_i}} - \frac{\hat{\lambda}\,\phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}{\Phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}\widehat{b'_{y_i}} - \frac{1}{\hat{\sigma}^2}\frac{\widehat{b_{y_i}}}{\widehat{b'_{y_i}}}\right\} = 0,$$

$$\sum_{i=1}^{n}\left\{(\hat{\nu} + \widehat{b_{y_i}})\left(\frac{y_i-\hat{\gamma}}{\hat{\sigma}}\right)\widehat{b'_{y_i}} - \hat{\lambda}\left(\frac{y_i-\hat{\gamma}}{\hat{\sigma}}\right)\frac{\phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}{\Phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}\widehat{b'_{y_i}} - \frac{1}{\hat{\sigma}^2}\left(\frac{y_i-\hat{\gamma}}{\hat{\sigma}}\right)\frac{\widehat{b_{y_i}}}{\widehat{b'_{y_i}}} - \frac{1}{\hat{\sigma}}\right\} = 0,$$

$$\sum_{i=1}^{n}\left\{\hat{\lambda}\,\frac{\phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}{\Phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]} - (\hat{\nu} + \widehat{b_{y_i}})\right\} = 0,$$

And

$$\sum_{i=1}^{n}\left\{\frac{\phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}{\Phi[\hat{\lambda}(\hat{\nu}+\widehat{b_{y_i}})]}(\hat{\nu} + \widehat{b_{y_i}})\right\} = 0.$$

Where $\widehat{b_{y_i}} = 2\, sinh([y_i-\hat{\gamma}]/\hat{\sigma})/\hat{\alpha}$, and $\widehat{b'_{y_i}} = 2\, cosh([y_i-\hat{\gamma}]/\hat{\sigma})/\hat{\alpha}\,\hat{\sigma}$.

The ML estimators for $\alpha, \gamma, \sigma, \nu$ and λ, are obtained by solving the likelihood equations which constitute a nonlinear system of equations, unfortunately there is no explicit solution, so that a computer program should be used to carry out a numerical procedure to solve this system of equations. In the next section, we will present the results of the ML estimator of the SSN distribution for specific values of the parameters $\alpha, \gamma, \sigma, \nu$ and λ, based on random samples generated by the proposed random number generator.

6. Numerical illustration

In this section, we present the results of Pearson's Chi-square test for the proposed random number generator, and the results of the ML estimator of the skewed sinh-normal distribution for specific values of the parameters $\alpha, \gamma, \sigma, \nu$ and λ. First, in order to study the effectiveness of the proposed random number generator, we will generate 100 sample of size 50 from the skewed sinh-normal distribution with parameters $\alpha, \gamma, \sigma, \nu$ and λ using the proposed random number generator, and then a pearson's chi-square test is performed for each sample, to examine how well those samples fit into the model at specific significance levels. In table 3, we list the χ^2 calculated in a pearson's chi-square test for 100 random samples of size 50 generated using our random number generator for $\alpha = 1.5, \gamma = 2, \sigma = 2, \nu = 1.5$ and $\lambda = 1$, along with the χ^2 tabulated at significance levels 0.01, 0.05, and 0.1. Table 3, show that at significance levels 0.01, 0.05, and 0.1 there are only 3, 6, and 10 random samples rejected out from the whole 100 random sample, respectively, which give us a good impression about the effectiveness of the proposed random number generator.

Second, we will generate 50 sample of sizes 20, 30, and 50 from the skewed sinh-normal distribution for different values of the parameters $\alpha, \gamma, \sigma, \nu$ and λ, using the proposed random number generator, and then the ML estimates for each sample will be obtained, along with the mean, standard deviation, mean square error, skewness, kurtosis, and distribution type of those estimates for the different sample sizes (for example, example see table 4). The steps of this procedure will be as the following:

1) Set the values of the parameters $\alpha, \gamma, \sigma, \nu$ and λ.
2) Generate 50 samples of sizes 20, 30, and 50 from $\mathrm{SSN}(\alpha, \gamma, \sigma, \nu, \lambda)$.
3) Obtain the ML estimates for $\alpha, \gamma, \sigma, \nu$ and λ, for different sample sizes.
4) Obtain the mean, standard deviation, mean square error, skewness, kurtosis, and distribution type for the ML estimates for different sample size.
5) Repeat 1:4 for different values of $\alpha, \gamma, \sigma, \nu$ and λ.

Table 3: Pearson's Chi Square Test for 100 Samples of Size 50 Generated from the SSN Distribution for $\alpha = 1.5, \gamma = 2, \sigma = 2, \nu = 1$, and $\lambda = 1$.

Sample No.	χ^2 calculated	Sample No.	χ^2 calculated	Sample No.	χ^2 calculated
1	0.506	35	9.842 ***	69	2.342
2	2.676	36	0.27	70	3.119
3	1.937	37	0.812	71	1.572
4	3.871	38	1.122	72	1.257
5	1.122	39	2.688	73	0.478
6	0.364	40	1.426	74	3.057
7	3.119	41	0.4	75	1.521
8	0.736	42	0.27	76	0.376
9	0.074	43	1.092	77	0.24
10	0.509	44	0.647	78	1.937
11	2.192	45	5.455 *	79	4.279
12	0.655	46	4.953 *	80	1.257
13	4.414	47	1.2	81	1.326
14	0.309	48	0.074	82	1.521
15	3.129	49	1.211	83	0.488
16	1.28	50	0.074	84	2.003
17	1.679	51	1.412	85	1.551
18	1.352	52	0.488	86	0.209
19	5.41 *	53	1.063	87	9.437 ***
20	2.558	54	0.71	88	0.569
21	0.108	55	1.332	89	0.48
22	4.388	56	0.036	90	2.788
23	8.021 **	57	6.984 **	91	3.083
24	5.093 *	58	3.459	92	2.593
25	3.638	59	4.174	93	0.24
26	2.106	60	2.333	94	0.634
27	0.783	61	2.593	95	4.478
28	0.392	62	0.993	96	10.433 ***
29	1.24	63	1.136	97	0.233
30	1.426	64	0.27	98	1.426
31	1.178	65	2.738	99	0.713
32	0.316	66	0.337	100	1.042
33	6.226 **	67	3.88		
34	0.448	68	0.584		

* Rejected at significance level 0.1, ** Rejected at significance levels 0.05, *** Rejected at significance level 0.01.

In table 4, we list the mean, standard deviation, mean square error, skewness, kurtosis, and distribution type for the ML estimates of the parameters $\alpha, \gamma, \sigma, \nu$ and λ, for 50 random samples of sizes 20,30 and 50 generated by the proposed random number generator for different values of the parameters. Table 4 shows that in general the standard deviation and the mean square error for the estimates of the parameters $\alpha, \gamma, \sigma, \nu$ and λ Decrease as the sample size increase, nevertheless, the standard deviation and the mean square error for the estimate of the scale parameter σ is often larger when it compared with the standard deviation and the mean square error for the estimates of the parameters α, γ, ν and λ. The distribution type for the parameters $\alpha, \gamma, \sigma, \nu$ and λ is vary between Type I, Type IV, and Type VI, while the Type I is the most common.

Table 4: The Mean, Standard Deviation, Mean Square Error, Skewness, Kurtosis, and Distribution Type for the Maximum Likelihood Estimates of the SSN Distribution.

Sample size		$\alpha = 1$	$\gamma = 2$	$\sigma = 2$	$\nu = 1.5$	$\lambda = 1$
		$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\nu}$	$\hat{\lambda}$
20	Mean	1.762	2.023	2.02	1.317	1.183
	Std	1.153	1.262	1.84	1.216	0.816
	MSE	1.884	1.561	3.318	1.482	0.686
	Skewness	2.14	2.349	2.31	2.157	0.603
	Kurtosis	11.606	8.422	7.691	8.333	3.213
	K	2.1519	-1.8393	-1.5597	-2.3427	-0.4485
	Dist. Type	Type VI	Type I	Type I	Type I	Type I
30	Mean	1.635	1.972	1.96	1.395	1.112
	Std	0.867	0.663	1.556	0.773	0.645
	MSE	1.14	0.432	2.376	0.597	0.42
	Skewness	0.723	1.035	2.686	0.846	-0.043
	Kurtosis	4.772	3.918	11.091	3.519	2.376
	K	0.2277	-0.7463	-2.8868	-0.5745	-0.0011
	Dist. Type	Type IV	Type I	Type I	Type I	Type I
50	Mean	1.496	1.918	1.647	1.359	1.001
	Std	0.609	0.615	0.637	0.683	0.569
	MSE	0.609	0.378	0.523	0.478	0.317
	Skewness	0.568	1.453	1.329	1.309	0.586
	Kurtosis	2.651	4.826	5.126	4.885	5.106
	K	-0.1605	-0.9297	-1.8311	-1.3484	0.0912
	Dist. Type	Type I	Type I	Type I	Type I	Type IV
		$\alpha = 1$	$\gamma = 0$	$\sigma = 2$	$\nu = 0.5$	$\lambda = 1$
		$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\nu}$	$\hat{\lambda}$
20	Mean	1.436	0.354	3.019	0.662	0.849
	Std	1.133	0.822	2.519	0.869	0.955
	MSE	1.448	0.787	7.259	0.765	0.9
	Skewness	0.585	1.402	0.959	1.439	0.485
	Kurtosis	2.205	5.987	2.671	5.302	2.222
	K	-0.1137	28.0263	-0.273	-1.4817	-0.0866
	Dist. Type	Type I	Type VI	Type I	Type I	Type I
30	Mean	1.273	-0.088	2.789	0.661	0.915
	Std	0.963	0.862	1.985	0.797	0.808
	MSE	0.984	0.736	4.485	0.679	0.663
	Skewness	1.119	-2.814	0.953	-1.556	0.7
	Kurtosis	4.4	13.351	2.614	9.105	2.453
	K	-1.2935	-5.8512	-0.2648	0.6145	-0.1705
	Dist. Type	Type I	Type I	Type I	Type IV	Type I
50	Mean	1.158	0.005	2.106	0.421	0.924
	Std	0.519	0.565	0.955	0.645	0.748
	MSE	0.289	0.313	0.905	0.414	0.554
	Skewness	0.319	0.16	1.082	0.983	0.705
	Kurtosis	2.293	6.382	3.447	5.115	3.466
	K	-0.0467	0.0033	-0.4522	0.6814	-0.7503
	Dist. Type	Type I	Type IV	Type I	Type IV	Type I

References

[1] Ashour, S.K., Abdel-hameed, M.A. (2010). Approximate skew normal distribution. Journal of Advanced Research 1:341–350. http://dx.doi.org/10.1016/j.jare.2010.06.004.

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

[3] Birnbaum, Z.W., Saunders, S.C. (1969). A new family of life distributions. Journal of Applied probability 6:319–327. http://dx.doi.org/10.2307/3212003.

[4] Galea, M., Leiva, V., Paula, G. A. (2004). Influence diagnostics in log Birnbaum–Saunders regression models. Journal of Applied Statistics 31:1049–1064. http://dx.doi.org/10.1080/0266476042000280409.

[5] Henze, N. (1986). A probabilistic representation of the skew–normal distribution. Scandinavian Journal of Statistics 13:271–275.

[6] Hoyt, J.P. (1968) a simple approximation to the standard normal probability density function. American Statistician.22, No.2, 25-56. http://dx.doi.org/10.2307/2681984.

[7] Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. Biometrika 36:149–176. http://dx.doi.org/10.1093/biomet/36.1-2.149.

[8] Johnson, N. L., Kotz, S., Balakrishnan, N. (1995). Continuous Univeriate Distributions –Volume 2. New York: Wiley.

[9] Leiva, V., Barros, M., Paula, G. A., Galea, M. (2007). Influence diagnostics in log - Birnbaum–Saunders regression models with censored data. Computational Statistics and Data Analysis 51:5694–5707. http://dx.doi.org/10.1016/j.csda.2006.09.020.

[10] Leiva, V., Vilca, F., Balakrishnan, N., Saunhueza, A. (2010). A skewed sinh–normal distribution and its properties and application to air pollution. Communication in Statistics – Theory and Methods 39:426–443. http://dx.doi.org/10.1080/03610920903140171.

[11] Owen, D. B. (1956). Tables for computing bivariate normal probabilities. Annals of Mathematical Statistics. 27, 1075–1090. http://dx.doi.org/10.1214/aoms/1177728074.

[12] Rieck, J. R. (1989). Statistical Analysis for the Birnbaum–Saunders Fatigue Life Distribution. Unpublished Ph.D. thesis, Department of Mathematical Sciences, Clemson University, USA.

[13] Rieck,J.R., Nedelman, J. R. (1991). A log-linear model for the Birnbaum–Saunders distribution. Technometrics 33:51–60. http://dx.doi.org/10.2307/1269007.

[14] Saunhueza, A., Leiva, V., Balakrishnan, N. (2008). The generalized Birnbaum–Saunders distribution and its theory, methodology and application. Communication in Statistics – Theory and Methods 37:645–670. http://dx.doi.org/10.1080/03610920701541174.