Sains Malaysiana 43(11)(2014): 1801–1809

An Empirical Assessment of the Closeness of Hidden Truncation and Additive Component based Skewed Distributions

(Penilaian Empirik Keakraban Pemangkasan Tersembunyi dan Komponen Tambahan berasaskan Taburan Terpencong)

PARTHA JYOTI HAZARIKA & SUBRATA CHAKRABORTY*

ABSTRACT

Hidden truncation (HT) and additive component (AC) are two well known paradigms of generating skewed distributions from known symmetric distribution. In case of normal distribution it has been known that both the above paradigms lead to Azzalini's (1985) skew normal distribution. While the HT directly gives the Azzalini's (1985) skew normal distribution, the one generated by AC also leads to the same distribution under a re-parameterization proposed by Arnold and Gomez (2009). But no such re-parameterization which leads to exactly the same distribution by these two paradigms has so far been suggested for the skewed distributions generated from symmetric logistic and Laplace distributions. In this article, an attempt has been made to investigate numerically as well as statistically the closeness of skew distributions generated by HT and AC methods under the same re-parameterization of Arnold and Gomez (2009) in the case of logistic and Laplace distributions.

Keywords: KS test; Kullback–Leibler (KL) distance; Monte Carlo integration; simulation; skew Laplace distribution; skew logistic distribution

ABSTRAK

Pemangkasan tersembunyi (HT) dan komponen tambahan (AC) adalah dua paradigma yang terkenal dalam menghasilkan taburan terpencong daripada taburan simetri. Dalam taburan normal ia telah diketahui bahawa kedua-dua paradigma di atas membawa terus kepada taburan pencongan normal (Azzalini 1985). Manakala HT terus memberikan taburan pencongan normal (Azzalini 1985), yang dijana oleh AC juga membawa kepada taburan yang sama di bawah pemparameteran semula yang dicadangkan oleh Arnold dan Gomez (2009). Tetapi tiada pemparameteran semula yang membawa kepada taburan yang sama oleh kedua-dua paradigma ini disarankan untuk taburan pencongan yang dihasilkan daripada simetri logistik dan taburan Laplace. Dalam artikel ini, usaha telah dibuat untuk mengkaji secara berangka dan statistik keakraban taburan pencongan yang dijana oleh kaedah HT dan AC di bawah pemparameteran semula Arnold dan Gomez (2009) bagi kes logistik dan taburan Laplace.

Kata kunci: Integrasi Monte Carlo; jarak Kullback-Leibler (KL); simulasi; taburan terpencong Laplace; taburan terpencong logistik; ujian KS

INTRODUCTION

The path breaking skew-normal distribution was first introduced by Azzalini (1985). A random variable Z is said to follow the skew normal distribution, SN(λ) if its probability density function (pdf) is given by,

$$f(z, \lambda) = 2\varphi(z)\Phi(\lambda z); -\infty < z < \infty, \lambda \in R,$$

where, ϕ and Φ are the pdf and cumulative distribution function (cdf) of the standard normal distribution, respectively. For $\lambda = 0$, SN (λ) reduces to standard normal distribution. Following Azzalini's (1985) seminal paper, lots of research work have so far been carried out to present different skew normal distributions derived from the underlying symmetric one to model asymmetric behavior of empirical data suitable under different situations (for a complete survey on univariate skew normal distributions see Chakraborty & Hazarika 2011). Besides skew normal distribution, skewed distribution based on other symmetric distributions, of which logistic, Laplace are notable have also been investigated by different authors (Kotz et al. 2001; Nadarajah 2009; Nekoukhou & Alamatsaz 2012; Wahed & Ali 2001). Huang and Chen (2007) proposed the general formula $f_z(z) = 2h(z)G(z)$, $z \in R$ for the construction of skew-symmetric distributions, where $h(.)$ is the pdf of a symmetric (about 0) distribution and the function $G(.)$, referred to as the skew function is a Lebesgue measurable function such that, $0 \le G(z) \le 1$ and $G(z)+G(-z) = 1$, $z \in R$, almost everywhere.

Following Huang and Chen (2007)'s method, Chakraborty et al. (2012, 2014a, and 2014b) introduced new skewed distributions based on logistic, normal and Laplace distribution which are suitable for modeling in multimodal data in the presence of skewness.

Arnold and Gomez (2009) pointed out that the skew distribution can be derived by hidden truncation (HT) and additive component (AC) methods and discussed their various aspects.

HIDDEN TRUNCATION (HT) METHOD

Let (X, Y) be a bivariate random variable with mean vector (μ_1, μ_2) and variance-covariance matrix $\sum = \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{2,2} \end{pmatrix}$. Now according to the HT method the distribution of a new random variable Z has been defined as the distribution of X conditional on $Y > a, a \in R$. The cdf of Z has been given by,

$$P(Z \leq z) = P(X \leq x | Y > a) = \frac{P(X \leq x, Y > a)}{P(Y > a)}. \quad (1)$$

This type of situation may arise in many real life applications (Arnold et al. 1993). For example: In admission to the programmes of a University/Colleges, usually marks obtained (X) in the entrance examination for the admission to a given programme of only those candidates whose marks (Y) in the qualifying examination exceed a given cut of marks (a) are considered for preparation of the final selection list for the admission; In the recruitment of police personals, the weight and/ or measure of chest (X) of only those candidates whose height (Y) is more than say 'a' are considered for preparation of the list of physically fit personals.

Alternatively, the probability distribution in (1) can be obtained as:

Let Y and W are two independent random variables having pdf ψ_1 and ψ_2 and cumulative distribution function (cdf) Ψ_1 and Ψ_2, respectively. Then according to the HT method the distribution of a new random variable Z has been defined as the conditional distribution of Y given the event $\{\lambda_0 + \lambda_1 Y > W\}$, $\lambda_0, \lambda_1 \in R$. The cdf of Z has been given by,

$$P(Z \leq z) = P(Y \leq z | \lambda_0 + \lambda_1 Y > W) = \frac{\int_{-\infty}^{z} \psi_1(y) \Psi_2(\lambda_0 + \lambda_1 y) dy}{P(\lambda_0 + \lambda_1 Y > W)}. \quad (2)$$

Differentiating both sides of (2) with respect to z, the pdf of Z has been obtained as,

$$f_{HT}(z; \lambda_0, \lambda_1) = \frac{\psi_1(y) \Psi_2(\lambda_0 + \lambda_1 y)}{P(\lambda_0 + \lambda_1 Y > W)}. \quad (3)$$

(For details see Azzalini 1986; Arnold et al. 1993; Arnold & Beaver 2000a)

ADDITIVE COMPONENT (AC) METHOD

Let U_1 and U_2 are two independent random variables having pdf (cdf), respectively, are $\psi_1(\Psi_1)$ and $\psi_2(\Psi_2)$. Now, define another r.v. $U_2(c)$, to be the r.v. U_2 truncated above c, $c \in R$ with density function,

$$f_{U_2(c)}(v) = \frac{\psi_2(v) I(v \leq c)}{\Psi_2(c)}. \quad (4)$$

where $I(.)$ is the usual indicator function. Then the AC method has been defined as the distribution of $U = U_1 + \delta U_2(c)$, $\delta \in R$. The resulting pdf of U has been given by

$$f_{AC}(u; \delta, c) = \int_{-\infty}^{c} \psi_1(u - \delta v) \frac{\psi_2(v)}{\Psi_2(c)} dv. \quad (5)$$

SKEW NORMAL DISTRIBUTION BASED ON HT AND AC METHOD

Based on HT In case of normal distribution the pdf of the skew normal distribution under HT has been given by,

$$f_{HT}(z; \mu, \sigma, \lambda_0, \lambda_1) = \left(\frac{z-\mu}{\sigma}\right) \Phi\left(\lambda_0 + \lambda_1 \frac{z-\mu}{\sigma}\right) \Big/ \sigma\, \Phi\left(\frac{\lambda_0}{\sqrt{1+\lambda_1^2}}\right), \quad (6)$$

where μ and σ are location and scale parameter, respectively. This is nothing but the two parameter skew normal distribution of Azzalini (1985). In particular, for $\lambda_0 = 0$, the distribution in (6) reduces to the skew normal distribution of Azzalini (1985). The multivariate extension of the above distribution has been studied by Arnold and Beaver (2000b, 2002).

Based on AC Considering ϑ and τ as location and scale parameter, respectively, the location scale extension of normal AC distribution has been given by,

$$f_{AC}(u; \vartheta, \tau, \delta, c) = \frac{1}{\tau\sqrt{1+\delta^2}} \phi\left(\frac{(u - \vartheta/\tau)}{\sqrt{1+\delta^2}} - \frac{\delta((u-\vartheta)/\tau)}{\sqrt{1+\delta^2}}\right) \Big/ \Phi(c). \quad (7)$$

Arnold and Gomez (2009) have shown that by using the re-parameterization

$$\mu = \vartheta, \sigma = \tau\sqrt{1+\delta^2},\ \lambda_0 = c\sqrt{1+\delta^2},\ \lambda_1 = -\delta$$

$$\text{or,}\quad \vartheta = \mu,\ \tau = \frac{\sigma}{\sqrt{1+\delta^2}}, c = \frac{\lambda_0}{\sqrt{1+\delta^2}},\ \delta = -\lambda_1 \quad (8)$$

the pdf derived by HT and AC method given in (6) and (7) leads to the same skew normal distribution.

SKEW LOGISTIC DISTRIBUTIONS BASED ON HT AND AC METHOD

Based on HT Let V_1 and V_2 are two i.i.d. standard Logistic random variables, $L(0,1)$, then the pdf in (3) becomes,

$$f_{HT}(z;\lambda_0,\lambda_1)=\frac{\psi_1(z)\Psi_2(\lambda_0+\lambda_1 z)}{P(\lambda_0+\lambda_1 V_1>V_2)}$$

where $\psi_1(.)$ and $\Psi_2(.)$ are the pdf and cdf of $L(0,1)$, respectively. Since $\psi_1(z) = \exp(z)(1 + \exp(z))^2$ and $\Psi_2(\lambda_0 + \lambda_1 z) = \exp(\lambda_0 + \lambda_1 z)/\{1 + \exp(\lambda_0 + \lambda_1 z)\}$, therefore the above pdf can be rewritten as (Arnold & Beaver 2000b; Arnold & Gomez 2009)

$$f_{HT}(z;\lambda_0,\lambda_1)=\frac{1}{P(\lambda_0+\lambda_1 V_1>V_2)}\frac{\exp(z)}{(1+\exp(z))^2}\frac{\exp(\lambda_0+\lambda_1 z)}{1+\exp(\lambda_0+\lambda_1 z)}. \tag{9}$$

Arnold and Gomez (2009) failed to give any compact form of $P(\lambda_0 + \lambda_1 V_1 > V_2)$ and prescribed that the constant must be evaluated numerically.

Based on AC Suppose that U_1 and U_2 are two independent standard logistic random variables then the corresponding pdf of Logistic additive component model has been obtained as,

$$f_{AC}(u;\delta,c)=\int_{-\infty}^{c}\psi(u-\delta v)\frac{\psi(v)}{\Psi(c)}dv,$$

where $\psi(.)$ and $\Psi(.)$ are the pdf and cdf of the standard Logistic distribution. Thus,

$$f_{AC}(u;\delta,c)=\frac{1+\exp(c)}{\exp(c)}\int_{-\infty}^{c}\frac{\exp(u-\delta v)}{(1+\exp(u-\delta v))^2}\frac{\exp(v)}{(1+\exp(v))^2}dv. \tag{10}$$

After introducing scale parameter τ and location parameter $\vartheta = 0$, the pdf becomes

$$f_{AC}(u;\delta,c)=\frac{1+\exp(c)}{\tau\exp(c)}\int_{-\infty}^{c}\frac{\exp\left(\frac{u}{\tau}-\delta v\right)}{\left\{1+\exp\left(\frac{u}{\tau}-\delta v\right)\right\}^2}\frac{\exp(v)}{(1+\exp(v))^2}dv. \tag{11}$$

The analytic form of the above density has not been available.

SKEW LAPLACE DISTRIBUTION BASED ON HIDDEN TRUNCATION

Based on HT Arnold and Gomez (2009) introduced the skew Laplace distribution based on HT model as follows:

Consider the pdf and cdf of standard Laplace distribution, respectively, given by

$$\psi_1(z) = \psi_2(z) = \psi(z) = \exp(-|z|)/2, -\infty < z < \infty \text{ and}$$

$$\Psi_1(z) = \Psi_2(z) = \Psi(z) = \begin{cases}\exp(z)/2, & z<0\\ 1-\{\exp(-z)\}/2, & z\geq 0\end{cases}$$

Then the pdf according to HT formulation has been given by,

$$f_{HT}(z;\lambda_0,\lambda_1) = \psi(z)\Psi(\lambda_0+\lambda_1 z) / P(\lambda_0 + \lambda_1 V_1 > V_2),$$

where, V_1 and V_2 are i.i.d. standard Laplace distribution

$$=\frac{\frac{1}{2}\exp(-|z|)\left\{\left[\frac{1}{2}\exp(\lambda_0+\lambda_1 z)I(\lambda_0+\lambda_1 z<0)\right]+\left[\left(1-\frac{1}{2}\exp(-\lambda_0-\lambda_1 z)I(\lambda_0+\lambda_1 z\geq 0)\right)\right]\right\}}{P(\lambda_0+\lambda_1 V_1>V_2)} \tag{12}$$

Here also the analytical form of $P(\lambda_0 + \lambda_1 V_1 > V_2)$ has been not available and has to be evaluated numerically.

Based on AC If U_1 and U_2 are two independent standard Laplace random variables having pdf, $\psi(u) = \frac{1}{2}\exp(-|u|)$, $-\infty < u < \infty$ then the pdf of skew Laplace based on AC model has been given by,

$$f_{AC}(u;\delta,c)=\int_{\infty}^{c}\psi(u-\delta v)\psi(v)/\Psi(c)dv,$$

where $\psi(.)$ and $\Psi(.)$ are the pdf and cdf of the standard Laplace distribution.

After introducing scale parameter τ and location parameter $\vartheta = 0$, the pdf becomes

$$f_{AC}(u;\tau,\delta,c)=\frac{1}{\tau}\int_{-\infty}^{c}\psi\left(\frac{u}{\tau}-\delta v\right)\frac{\psi(v)}{\Psi(c)}dv$$
$$=\frac{1}{4\tau\Psi(c)}\int_{-\infty}^{c}\exp(-|(u/\tau)-\delta v|)\exp(-|v|)dv. \tag{13}$$

Here as well, the density has not been available in analytic form.

Arnold and Gomez (2009) used the re-parameterization in (8) with Logistic and Laplace distribution and have graphically shown that the pdfs generated by the HT and the AC method do show some closeness, but they do not coincide as in the case of normal distribution.

The main objective of the present article was to test the statistically using KS test and numerically with KL Distance, how close are the skew distributions generated by HT and AC method in the case of Logistic and Laplace distributions under the re-parameterization of Arnold and Gomez (2009) (For some similar works on closeness and discrimination between two distributions see Gupta and Kundu (2003a, 2003b, 2004) and Pakyari (2011).

In the next section of this paper KS test has been used to check the closeness statistically while in the following section, numerical closeness has been investigated using KL distance.

STATISTICAL TEST TO VERIFY THE AGREEMENT OF HT AND AC BASED MODELS: KS TEST

Here, for a given set of parameters rejection sampling method has been used to generate 100 replication of random samples of size 1000 from the skew distribution using HT and AC method based on Logistic distribution and KS test has been performed the for their equality, similarly for the Laplace distribution. Tables 1 and 2 show the sample descriptive measures under HT and AC model based on Logistic and Laplace distribution, respectively, along with the percentages of agreement of HT and AC model based on the basis of KS test.

From Table 1 it can be seen that proportion of agreement between skew logistic distribution based on HT and AC model has been quite high ranging from 91 to 97%.

From Table 2, it may be concluded that proportion of agreement between skew Laplace distribution based on HT and AC has not been uniformly high as in the case of Logistic distribution. Here the proportion ranges from a low of 59% to a high of 98%.

NUMERICAL CLOSENESS OF HT AND AC BASED MODELS: KL DISTANCE

Kullback–Leibler (KL) distance has been one of the most widely used measures of the distance between two distributions. If f and g be two completely known continuous pdfs then $I(f,g)=\int_{-\infty}^{\infty} f(z)\log\{f(z)\}dz$ gives the information lost when g has been used to approximate f. This integral also gives the distance between g and f (Burnham & Anderson 2002). When the densities f and g are complicated, Monte Carlo integration (Robert & Casella 2004) can be employed to obtain close approximation to $I(f, g)$. In this section, the KL distance has been used to quantify the distance between HT (for fixed values of the parameters λ_0 and λ_1) and AC (with parameters determined through the re-parameterization given in (8)) models based on Logistic and Laplace as component distributions.

TABLE 1. Simulation results of HT and AC model based on logistic distribution

$\lambda_0 \downarrow$	$\lambda_1 \rightarrow$	0		1		2		3	
		AC	HT	AC	HT	AC	HT	AC	HT
0	Mean	-0.0042	-0.0010	0.0098	-0.0166	1.2365	1.2278	1.3127	1.3007
	Median	-0.0024	-0.0049	0.0089	-0.0166	1.0818	1.0510	1.1097	1.0833
	Mode	-0.0031	-0.0017	0.0090	-0.0163	1.1813	1.1614	1.2381	1.2197
	Skewness	-0.0065	0.0138	0.0073	0.0030	0.7286	0.9206	1.0483	1.1462
	Kurtosis	4.1130	4.1392	4.1745	4.1092	4.4062	4.9548	5.1200	5.4352
	KS	96%		91%		93%		97%	
1	Mean	-0.0063	-0.0041	0.6754	0.6872	0.9749	0.9865	1.1175	1.1084
	Median	-0.0120	-0.0022	0.6047	0.5713	0.8219	0.8082	0.9145	0.8897
	Mode	-0.0071	-0.0033	0.6516	0.6441	0.9198	0.9190	1.0432	1.0278
	Skewness	-0.0099	-0.0109	0.2466	0.5360	0.7185	0.9153	1.0056	1.1034
	Kurtosis	4.1048	4.1718	3.7855	4.1876	4.4376	4.9225	4.9214	5.2297
	KS	97%		93%		95%		97%	
2	Mean	0.0034	-0.0101	0.4345	0.4491	0.7598	0.7651	0.9366	0.9359
	Median	0.0045	-0.0137	0.3664	0.3452	0.6098	0.5907	0.7355	0.7217
	Mode	0.0028	-0.0104	0.4118	0.4095	0.7054	0.6993	0.8639	0.8569
	Skewness	0.0089	0.0057	0.2259	0.4837	0.6574	0.8477	0.9496	1.0518
	Kurtosis	4.1680	4.1226	3.7011	4.0751	4.1013	4.5754	4.7022	5.0216
	KS	97%		95%		96%		95%	
3	Mean	0.0022	-0.0112	0.2694	0.2783	0.5758	0.5846	0.7796	0.7787
	Median	-0.0011	-0.0107	0.2062	0.1935	0.4306	0.4193	0.5850	0.5740
	Mode	0.0012	-0.0108	0.2483	0.2446	0.5236	0.5219	0.7094	0.7020
	Skewness	0.0077	0.0063	0.2052	0.4217	0.6115	0.7927	0.8810	0.9903
	Kurtosis	4.2111	4.0962	3.6061	4.0114	3.9675	4.4523	4.4489	4.7925
	KS	94%		97%		95%		93%	

TABLE 2. Simulation results of HT and AC model based on Laplace distribution

$\lambda_0 \downarrow$	$\lambda_1 \rightarrow$	0		1		2		3	
		AC	HT	AC	HT	AC	HT	AC	HT
0	Mean	-0.0036	-0.0004	0.7064	0.7444	0.8927	0.8834	0.9480	0.9307
	Median	0.0012	-0.0025	0.6071	0.5287	0.7113	0.6446	0.7226	0.6728
	Mode	-0.0023	-0.0010	0.6773	0.6703	0.8287	0.7951	0.8650	0.8338
	Skewness	-0.0086	0.0219	0.3801	1.1242	1.0432	1.4568	1.4331	1.6321
	Kurtosis	5.8406	5.8816	4.9677	6.0765	5.6966	6.9102	6.7996	7.4259
	KS	96%		67%		80%		83%	
1	Mean	0.0081	-0.0114	0.4004	0.4377	0.6080	0.6332	0.7165	0.7271
	Median	0.0033	-0.0078	0.3022	0.2419	0.4306	0.3886	0.4944	0.4630
	Mode	0.0063	-0.0101	0.3714	0.3698	0.5448	0.5448	0.6339	0.6291
	Skewness	0.0397	-0.0133	0.3578	1.1097	1.0285	1.4837	1.4203	1.6612
	Kurtosis	5.9244	5.8870	4.8363	6.3526	5.7136	7.0565	6.7460	7.5621
	KS	98%		59%		83%		95%	
2	Mean	0.0010	-0.0055	0.2342	0.2536	0.4341	0.4550	0.5604	0.5723
	Median	-0.0001	-0.0064	0.1522	0.1119	0.2720	0.2346	0.3483	0.3199
	Mode	0.0002	-0.0053	0.2088	0.2009	0.3753	0.3735	0.4820	0.4781
	Skewness	0.0103	0.0128	0.3163	0.9251	0.9555	1.4070	1.3561	1.6114
	Kurtosis	5.9189	5.8845	4.6523	6.0677	5.5164	6.9701	6.6050	7.3901
	KS	94%		66%		79%		89%	
3	Mean	0.0018	-0.0076	0.1410	0.1355	0.3168	0.3297	0.4437	0.4549
	Median	-0.0015	-0.0044	0.0807	0.0505	0.1762	0.1449	0.2487	0.2231
	Mode	0.0011	-0.0073	0.1197	0.1008	0.2634	0.2592	0.3700	0.3678
	Skewness	0.0022	0.0301	0.2889	0.6589	0.8673	1.2433	1.2622	1.5219
	Kurtosis	5.9839	5.7938	4.4933	5.5562	5.2254	6.5177	6.2381	7.2272
	KS	93%		64%		75%		87%	

For Logistic distribution, the formula of KL distance between HT and AC model has been given by,

$$I(f_{HT}(z,\lambda_0,\lambda_1), f_{AC}(z;\tau,\delta,c)) =$$

$$\int_{-\infty}^{\infty}\left\{\frac{1}{P(\lambda_0+\lambda_1V_1>V_2)}\frac{\exp(z)}{(1+\exp(z))^2}\frac{\exp(\lambda_0+\lambda_1 z)}{1+\exp(\lambda_0+\lambda_1 z)}\right\}\times$$

$$\log\left[\frac{1}{P(\lambda_0+\lambda_1V_1>V_2)}\frac{\exp(z)}{(1+\exp(z))^2}\frac{\exp(\lambda_0+\lambda_1 z)}{1+\exp(\lambda_0+\lambda_1 z)}\times\right.$$

$$\left.\left[\frac{1+\exp(c)}{\tau\exp(c)}\int_{-\infty}^{c}\frac{\exp\left(\frac{z}{\tau}-\delta v\right)}{\left\{1+\exp\left(\frac{z}{\tau}-\delta v\right)\right\}^2}\frac{\exp(v)}{(1+\exp(v))^2}dv\right]^{-1}\right]dz \tag{14}$$

where the value of λ_0, λ_1, τ, δ and c are known. As the integrations are quite involved, the Monte Carlo Integration has been applied here to obtain close approximation to the result of the above integration. The Monte Carlo estimation of the above integral has been given by,

$$I(f_{HT}(z,\lambda_0,\lambda_1), f_{AC}(z;\tau,\delta,c)) =$$

$$\left[\sum_z\left(\frac{\frac{\exp(z)}{(1+\exp(z))^2}\frac{\exp(\lambda_0+\lambda_1 z)}{1+\exp(\lambda_0+\lambda_1 z)}}{P(\lambda_0+\lambda_1V_1>V_2)}\right)\right]^{-1}$$

$$\sum_z\left[\left\{\frac{\frac{\exp(z)}{(1+\exp(z))^2}\frac{\exp(\lambda_0+\lambda_1 z)}{1+\exp(\lambda_0+\lambda_1 z)}}{P(\lambda_0+\lambda_1V_1>V_2)}\right\}\times\right.$$

$$\log\left\{\frac{\frac{\exp(z)}{(1+\exp(z))^2}\frac{\exp(\lambda_0+\lambda_1 z)}{1+\exp(\lambda_0+\lambda_1 z)}}{P(\lambda_0+\lambda_1V_1>V_2)}\times\right.$$

$$\left[\frac{1}{\tau}\int_{-\infty}^{c}\frac{\exp\left(\frac{z}{\tau}-\delta v\right)}{\left(1+\exp\left(\frac{z}{\tau}-\delta v\right)\right)^{2}}\,\frac{\dfrac{\exp(v)}{\left(1+\exp(v)\right)^{2}}}{\dfrac{\exp(c)}{1+\exp(c)}}\,dv\right]^{-1}\Bigg]\Bigg], \qquad (15)$$

where the sum is taken over a large number of equally spaced values of Z within the range of Z.

In a similar manner for the Laplace distribution, the formula of KL distance between HT and AC model and its Monte Carlo estimator are, respectively, given by

$$I(f_{HT}(z,\lambda_0,\lambda_1), f_{AC}(z;\tau,\delta,c)) =$$

$$\int_{-\infty}^{\infty}\left\{\frac{\frac{1}{2}\exp(-|z|)\left\{\left[\frac{1}{2}\exp(\lambda_0+\lambda_1 z)I(\lambda_0+\lambda_1 z<0)\right]+\left[(1-\frac{1}{2}\exp(-\lambda_0-\lambda_1 z))I(\lambda_0+\lambda_1 z\geq 0)\right]\right\}}{P(\lambda_0+\lambda_1 V_1>V_2)}\right\}$$

$$\log\left[\left\{\frac{\frac{1}{2}\exp(-|z|)\left\{\left[\frac{1}{2}\exp(\lambda_0+\lambda_1 z)I(\lambda_0+\lambda_1 z<0)\right]+\left[(1-\frac{1}{2}\exp(-\lambda_0-\lambda_1 z))I(\lambda_0+\lambda_1 z\geq 0)\right]\right\}}{P(\lambda_0+\lambda_1 V_1>V_2)}\right\}\times\right.$$

$$\left.\left[\frac{1}{4\tau\,\Psi(c)}\int_{-\infty}^{c}\exp(-|z/\tau-\delta v|)\exp(-|v|)\,dv\right]^{-1}\right]dz \qquad (16)$$

and

$$\hat{I}(f_{HT}(z,\lambda_0,\lambda_1), f_{AC}(z;\tau,\delta,c)) =$$

$$\left\{\sum_{z}\left[\left\{\frac{\frac{1}{2}\exp(-|z|)\left\{\left[\frac{1}{2}\exp(\lambda_0+\lambda_1 z)I(\lambda_0+\lambda_1 z<0)\right]+\left[(1-\frac{1}{2}\exp(-\lambda_0-\lambda_1 z))I(\lambda_0+\lambda_1 z\geq 0)\right]\right\}}{P(\lambda_0+\lambda_1 V_1>V_2)}\right\}\times\right.\right.$$

$$\log\left[\left\{\frac{\frac{1}{2}\exp(-|z|)\left\{\left[\frac{1}{2}\exp(\lambda_0+\lambda_1 z)I(\lambda_0+\lambda_1 z<0)\right]+\left[(1-\frac{1}{2}\exp(-\lambda_0-\lambda_1 z))I(\lambda_0+\lambda_1 z\geq 0)\right]\right\}}{P(\lambda_0+\lambda_1 V_1>V_2)}\right\}\right.$$

$$\left.\left.\left.\left[\frac{1}{4\tau\,\Psi(c)}\int_{-\infty}^{c}\exp(-|z/\tau-\delta v|)\exp(-|v|)\,dv\right]^{-1}\right]\right\}\right/$$

$$\left\{\sum_{z}\left[\frac{\frac{1}{2}\exp(-|z|)\left\{\left[\frac{1}{2}\exp(\lambda_0+\lambda_1 z)I(\lambda_0+\lambda_1 z<0)\right]+\left[(1-\frac{1}{2}\exp(-\lambda_0-\lambda_1 z))I(\lambda_0+\lambda_1 z\geq 0)\right]\right\}}{P(\lambda_0+\lambda_1 V_1>V_2)}\right]\right\} \qquad (17)$$

TABLE 3. KL distance between HT and AC model based on

Parameters of HT		Parameters of AC Model after using transformation			KL distance between HT and AC model based on	
λ_0	λ_1	c	τ	δ	Logistic distribution	Laplace distribution
0	0	0	1	0	0.0000	0.0000
0	1	0	0.7071	-1	0.0042	0.0201
0	2	0	0.4472	-2	0.0013	0.0108
0	3	0	0.3162	-3	0.0004	0.0055
1	0	1	1	0	0.0000	0.0000
1	1	0.7071	0.7071	-1	0.0047	0.0195
1	2	0.4472	0.4472	-2	0.0017	0.0134
1	3	0.3162	0.3162	-3	0.0005	0.0075
2	0	2	1	0	0.0000	0.0000
2	1	1.4142	0.7071	-1	0.0043	0.0141
2	2	0.8944	0.4472	-2	0.0019	0.0119
2	3	0.6324	0.3162	-3	0.0007	0.0076
3	0	3	1	0	0.0000	0.0000
3	1	2.1213	0.7071	-1	0.0035	0.0127
3	2	1.3416	0.4472	-2	0.0018	0.0097
3	3	0.9486	0.3162	-3	0.0008	0.0065

where the value of λ_0, λ_1, τ, δ and c are known.

Table 3 presents the KL distance between HT and AC model based on logistic distribution and Laplace distribution. Figure 1(a) to 1(l) shows the density of HT (dotted line) and AC (line) based skew Logistic distributions while Figure 2(a) to 2(l) shows the density of HT (dotted line) and AC (red line) model based on Laplace distribution.

From Table 3 it can be seen that the KL distance in case of logistic distribution is very low with maximum being 0.0047 and the minimum is 0.00. But in the case of Laplace distribution, the KL distance has not been uniformly low with the maximum distance recorded being 0.0201 and the minimum being 0.00. Furthermore it has been observed that the distance is higher whenever $\lambda_1 = 1$.

The densities have been plotted to visually inspect their closeness. Figure 1(a) to 1(l) displays the skew logistic densities generated by HT and AC. Here it has been observed that the peak ness of the hidden truncation model as compared to AC model is high, which can also be verified from the values of the kurtosis presented in Table 1. Figure 2(a) to 2(l) which displays the skew Laplace densities also tells the same story. Since, it has been theoretically proven that for any values of λ_0 when

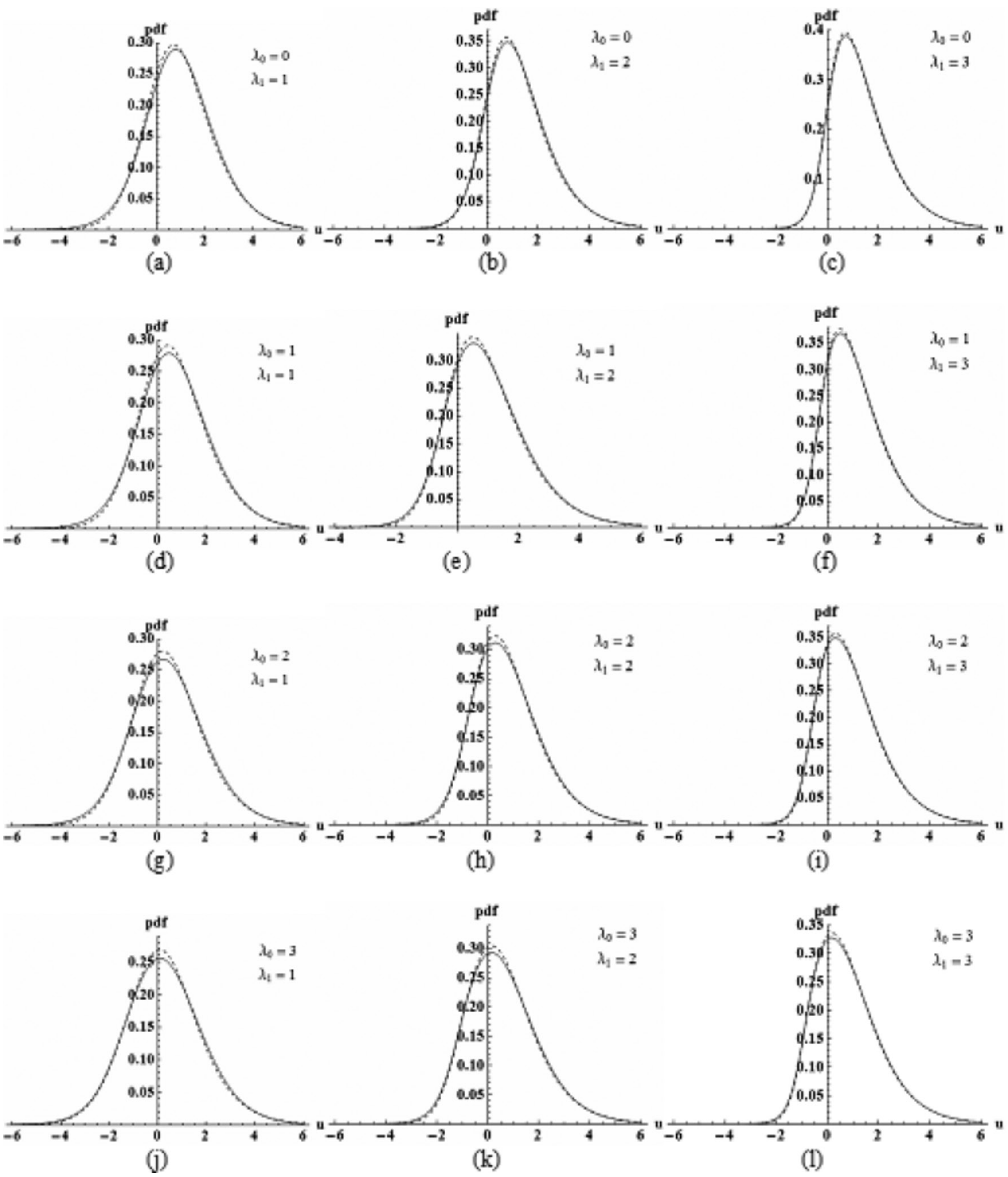

FIGURE 1. (a-l): Plots of densities of HT and AC model based on logistic distribution

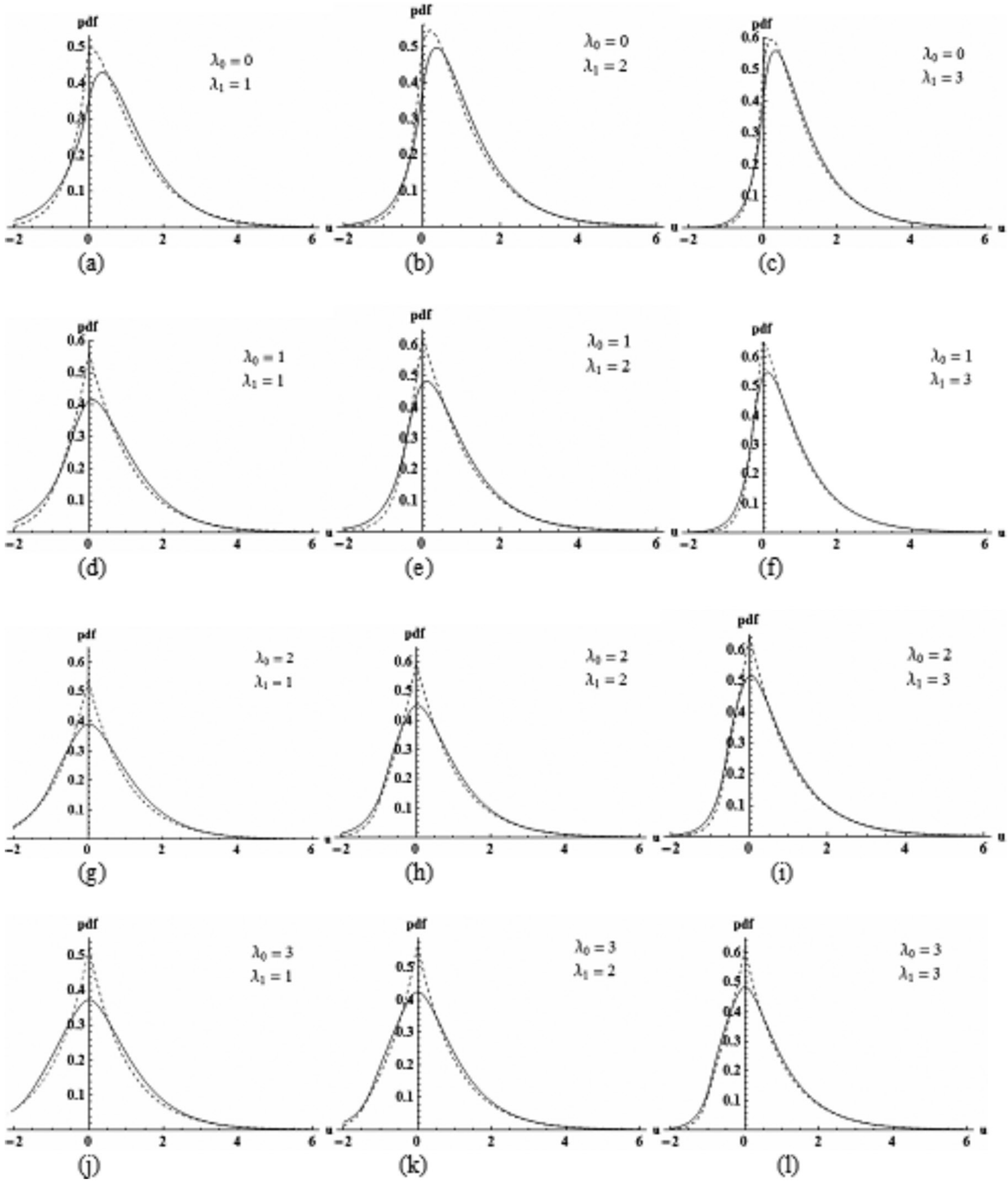

FIGURE 2. (a-l): Plots of densities of HT and AC model based on Laplace distribution

$\lambda_1 = 0$ both the density are the same, we have not shown the plots when $\lambda_1 = 0$ (see rows 1, 5, 9 and 13 of Table 3).

CONCLUSION AND COMMENTS

From the present investigation, it has been apparent that the skew distributions generated by HT and AC method when the component distributions are logistic and Laplace may not always be close to each other under the re-parameterization of Arnold and Gomez (2009). Further research will be needed to see whether some other transformation exist which will bring the skew models generated by the two paradigms closer.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the suggestions and comments of the anonymous referee which helped immensely to make substantial improvements of the paper.

REFERENCES

Arnold, B.C. & Beaver, R.J. 2002. Skewed multivariate models related to hidden truncation and/or selective reporting (with discussion). *Test* 11: 7-54.

Arnold, B.C. & Beaver, R.J. 2000a. The skew Cauchy distribution. *Statist. Prob. Lett.* 49: 285-290.

Arnold, B.C. & Beaver, R.J. 2000b. Hidden truncation models. *Sankhya* A62: 22-35.

Arnold, B.C., Beaver, R.J., Groeneveld, R.A. & Meeker, W.Q. 1993. The non-truncated marginal of a truncated bivariate normal distribution. *Psychometrika* 58: 471-478.

Arnold, B.C. & Gomez, H.W. 2009. Hidden truncation and additive components: Two alternative skewing paradigms. *Calcutta Statistical Association Bulletin* 61: 241-244.

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

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

Burnham, K.P. & Anderson, D.R. 2002. *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*. 2nd ed. New York: Springer.

Chakraborty, S. & Hazarika, P.J. 2011. A survey of the theoretical developments in univariate skew normal distributions. *Assam. Statist. Rev.* 25(1): 41-63.

Chakraborty, S., Hazarika, P.J. & Ali, M. Masoom. 2014a. A multimodal skew laplace distribution. *Pak. J. Statist***.** 30(2): 253-264**.**

Chakraborty, S., Hazarika, P.J. & Ali, M. Masoom. 2014b. A multimodal skewed extension of normal distribution: Its properties and applications. *Statistics: A Journal of Theoretical and Applied Statistics* DOI: 10.1080/02331888.2014.908880.

Chakraborty, S., Hazarika, P.J. & Ali, M. Masoom. 2012. A new skew logistic distribution and its properties. *Pak. J. Statist.* 28(4): 513-524.

Gupta, R.D. & Kundu, D. 2004. Discriminating between gamma and generalized exponential distributions. *J. Statist. Comput. Simul.* 74(2): 107-121.

Gupta, R.D. & Kundu, D. 2003a. Closeness of gamma and generalized exponential distribution. *Communications in Statistics-Theory and Methods* 32(4): 705-721.

Gupta, R.D. & Kundu, D. 2003b. Discriminating between Weibull and generalized exponential distributions. *Computational Statistics & Data analysis* 43(2): 179-196.

Huang, W.J. & Chen, Y.H. 2007. Generalized skew Cauchy distribution. *Statist. Probab. Lett.* 77: 1137-1147.

Kotz, S., Kozubowski, T.J. & Podgorski, K. 2001. *The Laplace Distribution and Generalizations*. Berlin: Birk Hauser.

Nadarajah, S. 2009. The skew logistic distribution. *Adv. Stat. Anal.* 93: 187-203.

Nekoukhou, V. & Alamatsaz, M.H. 2012. A family of skew-symmetric-Laplace distributions. *Statistical Papers* 53(3): 685-696.

Pakyari, R. 2011. Discriminating between generalized exponential, geometric extreme exponential and Weibull distributions. *J. Statist. Comput. Simul.* 80(12): 1403-1412.

Robert, C.P. & Casella, G. 2004. *Monte Carlo Statistical Methods*. 2nd ed. New York: Springer.

Wahed, A.S. & Ali, M. Masoom. 2001. The skew logistic distribution. *J. Stat. Res.* 35: 71-80.

Partha Jyoti Hazarika
Department of Statistics
North-Eastern Hill University
Shillong 793022, Meghalaya
India

Subrata Chakraborty*
Department of Statistics, Dibrugarh University
Dibrugarh 786004, Assam
India

*Corresponding author; email: subrata_arya@yahoo.co.in

Received: 26 June 2013
Accepted: 31 March 2014