

A NEW FAMILY OF SKEWED SLASH DISTRIBUTIONS GENERATED BY THE NORMAL KERNEL

B. Punathumparambath

1. INTRODUCTION

Symmetric distributions generalizing normality have got more attention in the statistical literature. One such extension is the class of standard slash (slash normal) distribution proposed by Kafadar (see, 1982 and 1988). The standard slash distribution is obtained as the distribution of the ratio, $X = \frac{Y}{U^{1/q}}$ where Y is a standard normal random variable; U is an independent uniform random variable over the interval $(0, 1)$ and $q > 0$. Slash distributions have heavier tails than those of the normal distributions. It generalizes the normal distribution by allowing a tail parameter which can adjust to absorb heavy tails. It has simple structure and is convenient to be generated. Thus it has appealing applications in simulating and fitting data with heavy tails. When $q = 1$ we obtain the canonical slash, which has the same tail heaviness as the Cauchy. The probability density function (*pdf*) of the univariate slash distribution for $q = 1$ is given by,

$$f(x,1) = \begin{cases} \frac{\phi(0) - \phi(x)}{x^2}; & x \neq 0 \\ \frac{\phi(0)}{2}; & x = 0 \end{cases} \quad (1)$$

where $\phi(\cdot)$ denotes the *pdf* of the standard normal distribution. As $q \rightarrow \infty$ the slash normal (slash) distribution yields the usual normal distribution. General properties of this distribution are studied in Rogers and Tukey (1972) and in Mosteller and Tukey (1977). Maximum likelihood estimates (MLEs) of the related location-scale family are discussed in Kadafar (1982). Recently, Wang and Genton (2006) described multivariate and skew-multivariate extensions of the slash distribution and investigated its properties. They applied their skew-slash distributions to fit AIS and glass-fiber data. In the present work we introduce a new family of distribution which is a generalization of the skew slash normal (skew-slash) distribution.

The error distribution in the regression models need not obey Gaussian law in all situations, for example, regression models applied in the field of biology, economics, psychology and sociology. Hence, there is a special interest in constructing distributions that could describe symmetry and skewness observed in the data. In the last two decades there has been substantial work in the area of skew-normal (SN) and related distributions. Gupta *et al.* (2002) provided detailed discussion of skew-symmetric models based on the normal, Students t, Cauchy, Laplace, logistic and uniform distributions. Nadarajah and Kotz (2003) developed various skewed models generated by the normal kernel: the skew normal-normal model, the skew normal-t model, the skew normal-Cauchy model, the skew normal-Laplace model, the skew normal-logistic model and the skew normal-uniform model. The present work is a generalization of the skewed distributions generated by the normal kernel. The construction of the above univariate skew-symmetric models is based on the general result due to Azzalini (1985) and is given below.

Lemma 1. Let X be a random variable with density function $f(y)$ symmetric about 0, and Z a random variable with absolutely continuous distribution function $G(y)$ such that $G'(y)$ is symmetric about 0, then for any $\lambda > 0$,

$$h(y/\lambda) = 2f(y)G(\lambda y), \quad -\infty < y < \infty \quad (2)$$

is density function for any real λ .

For example, the skew normal distribution is obtained by taking $f \equiv \phi$ and $G \equiv \Phi$ in (2).

This article is organized as follows. Section 2 defines the skew slash normal-normal distribution and discusses its various properties. Section 3, introduces the skew slash normal-Cauchy distribution and explores its properties. Section 4, introduces the skew slash normal-Laplace distribution and its various properties are studied.

2. SKEW SLASH NORMAL-NORMAL DISTRIBUTION

The skew slash normal-normal distribution can be defined as the distribution of the ratio $X = \frac{Y}{U^{1/q}}$ where Y is a skew normal-normal (SNN) random variable and U is an independent uniform random variable over $(0, 1)$ and $q > 0$. It is denoted by $X \sim \text{SSNN}(\gamma, \sigma, \lambda; q)$ and is a generalization of the slash distribution. The distribution of Y is obtained by taking f to be the pdf of a normal distribution with scale parameter γ and g to be a normal pdf with zero mean and variance σ^2 . Then from (2) we get the *pdf* of the skew normal-normal distribution as given below.

$$f(y; \gamma, \sigma, \lambda) = \frac{\sqrt{2q}}{\sqrt{\pi\gamma}} e^{\frac{-x^2}{2\gamma^2}} \Phi\left(\frac{\lambda y}{\sigma}\right), -\infty < y < \infty. \tag{3}$$

The pdf of skew slash normal-normal can be given by,

$$g(x; q) = \int_0^1 u^{1/q} f(xu^{1/q}) du, -\infty < x < \infty, \tag{4}$$

where $f(\cdot)$ is the skew normal-normal (SNN) distribution.

For the substitution $v = u^{1/q}$, we get the skew slash normal-normal distribution and is defined as follows.

Definition 1. A random variable X denoted by $X \sim SSNN(\gamma, \sigma, \lambda; q)$ is said to have a skew slash normal-normal distribution if its probability density function is

$$h(x; \gamma, \sigma, \lambda, q) = \frac{\sqrt{2q}}{\sqrt{\pi\gamma}} \int_0^1 v^q e^{\frac{-x^2 v^2}{2\gamma^2}} \Phi\left(\frac{\lambda xv}{\sigma}\right) dv, -\infty < x < \infty. \tag{5}$$

where $\gamma > 0, \sigma > 0, q > 0$ and $\lambda \in \mathbf{R}$.

The characteristic function of the SSNN distribution is given by,

$$\varphi_X(t) = \int_0^1 \Psi_Y(tu^{-1/q}) du$$

where $\Psi_Y(\cdot)$ is the characteristic function of the skew normal-normal distribution (Nadarajah and Kotz (2003)) and is given by

$$\Psi_Y(t) = E(e^{itY}) = 2 \exp\left(\frac{-\gamma^2 t^2}{2}\right) \Phi\left(\frac{it\lambda\gamma}{\sigma} / \sqrt{\frac{1}{\gamma^2} + \frac{\lambda^2}{\sigma^2}}\right)$$

The mean of X is given by

$$E(X) = \frac{q}{q-1} \frac{\sqrt{2\lambda\gamma^2}}{\sqrt{\pi\sigma^2 \left(1 + \frac{\gamma^2 \lambda^2}{\sigma^2}\right)}}, \quad q > 1$$

From the Figure 1, we can see that SSNN distribution has heavier tails, asymmetry of varying degrees and peakedness than the normal distribution. SSNN density is symmetric, negatively skewed and positively skewed for $\lambda=0, \lambda < 0, \lambda > 0$ respectively. The main feature of the slash skew normal-normal distribution in (5) is that the parameters λ and q control skewness, kurtosis and tail behaviour. The following properties of the probability density function (5) hold. For most of them the proof is immediate.

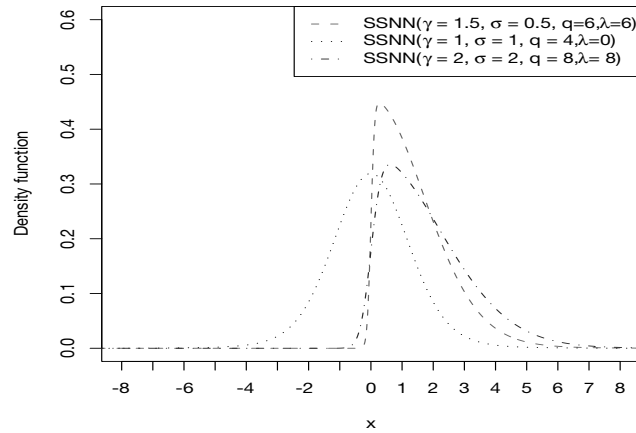


Figure 1 – Skew slash normal-normal density functions for various values of parameters.

Remark 1. The limiting distribution of the SSNN distribution is, as $q \rightarrow \infty$ the skew normal-normal distribution (SNN). The limiting distribution of SSNN as $q \rightarrow \infty$ and $\lambda = 0$ is univariate normal distribution.

Remark 2. For $\gamma = \sigma = 1$ we get the univariate skew slash distribution and for $\gamma = \sigma = 1$ and $\lambda = 0$, we get the slash distribution. Also the limiting distribution of SSNN as $q \rightarrow \infty$ and $\gamma = \sigma = 1$ is univariate skew normal distribution.

Remark 3. Note that the skew slash normal-normal random variable in (5) is a scale mixture of the skew normal-normal random variable and so it can be represented as,

$$X | (U = u) \sim \text{SNN}(0, u^{-1/q} \gamma, u^{-1/q} \sigma, \lambda) \text{ with } U \sim U(0, 1)$$

Remark 4. The density has point mass at $x = 0$, that is,

$$f(0; \gamma, \sigma, \lambda, q) = \frac{1}{\sqrt{2\pi}\gamma} \frac{q}{(q+1)}$$

Remark 5. If X has a skew slash normal-normal distribution with skew parameter λ , then $-X$ also has a skew slash normal-normal distribution with skewness parameter $-\lambda$. That is, if $X \sim \text{SSNN}(\gamma, \sigma, \lambda; q)$ then $-X \sim \text{SSNN}(\gamma, \sigma, -\lambda; q)$.

Remark 6. If X has a skew slash normal-normal distribution and a be a nonzero real constant, then aX also has a skew slash normal-normal distribution. That is, if $X \sim \text{SSNN}(\gamma, \sigma, \lambda; q)$ then $aX \sim \text{SSNN}(|a| \gamma, |a| \sigma, \lambda_a; q)$, where $\lambda_a = \text{sign}(a)\lambda$ and $\text{sign}(a) = 1$ for $a > 0$, -1 for $a < 0$, 0 for $a = 0$. This implies that the skew-slash normal-normal distribution is invariant under linear transformations.

Proposition 1. If $X_\lambda(y)$ and $X_0(y)$ are respectively the skew slash normal-normal distribution with skewness parameter λ and slash random variables. Then

$$X_\lambda(y) \rightarrow |X_0(y)| \text{ as } \lambda \rightarrow \infty$$

That is, when $\lambda \rightarrow \infty$ the skew slash normal-normal distribution tends to half slash distribution.

Proposition 2. Perturbation invariance: For any even function b , $b(X_\lambda) \stackrel{d}{=} b(X)$. Hence $X_\lambda^2 \stackrel{d}{=} X^2$, where X_λ follows a skew slash normal-normal and X is slash distribution.

Proposition 3. The skew slash normal-normal distribution is unimodal since the distribution monotonically increasing on $(-\infty, 0)$ and monotonically decreasing on $(0, \infty)$. The distribution have mode at $x = 0$ and the value at mode is

$$f_{\gamma, \lambda, \sigma} = \frac{q}{(q+1)} \frac{1}{\sqrt{2\pi\gamma}}$$

From remarks (1), (2) and (3) we can see that the skew slash normal-normal is a generalization of the normal, skew normal, skew normal-normal, slash and skew slash distributions. This distribution can be useful in analyzing data sets which are skewed, heavy tailed and peaked. This distribution is useful in simulation studies where it can introduce distributional challenges in order to evaluate a statistical procedure. It is also useful in analyzing data sets that do not follow the normal law. Additional flexibility can be introduced in the skew-slash normal-normal distribution by allowing higher order odd polynomials in the skewing function $\Phi(\cdot)$ in (5). For instance, an odd polynomial of order three would yield a distribution that can model bimodality, see Ma and Genton (2004) for further discussions on this topic.

3. SKEW SLASH NORMAL-CAUCHY DISTRIBUTION

The skew slash normal-Cauchy distribution can be defined as the distribution of the ratio $X = Y / U^{1/q}$ where Y is a skew normal-Cauchy random variable and U is an independent uniform random variable over $(0, 1)$ and $q > 0$. It is denoted by $X \sim SSNC(\gamma, \sigma, \lambda; q)$.

Definition 2. A random variable X denoted by $X \sim SSNC(\gamma, \sigma, \lambda; q)$ is said to have a skew slash normal-Cauchy distribution if its probability density function is

$$h(x; \gamma, \sigma, \lambda) = \frac{q\sqrt{2}}{\sqrt{\pi\gamma}} \int_0^1 v^q e^{-\frac{x^2 v^2}{2\gamma^2}} \left\{ 1 + \frac{2}{\pi} \arctan\left(\frac{\lambda x v}{\sigma}\right) \right\} dv, -\infty < x < \infty. \quad (6)$$

where $\gamma > 0, \sigma > 0, q > 0$ and $\lambda \in \mathbb{R}$.

The mean of X is given by

$$E(X) = \frac{q}{q-1} \frac{2\sqrt{2}\gamma}{\sqrt{\pi}} \exp\left(-\frac{\sigma^2}{2\gamma^2\lambda^2}\right) \Phi\left(-\frac{\sigma}{\gamma\lambda}\right), \quad q > 1$$

Remark 7. Note that the SSNC random variable in (6) is a scale mixture of the skew normal-Cauchy (SNC) random variable and so it can be represented as,

$$X | (U = u) \sim \text{SNC}(0, u^{-1/q} \gamma, u^{-1/q} \sigma, \lambda) \text{ with } U \sim U(0,1)$$

The probability density function (6) has the similar properties of the (5).

Figure 2 gives the probability density function of the skew slash normal-Cauchy density for various values of the parameters.

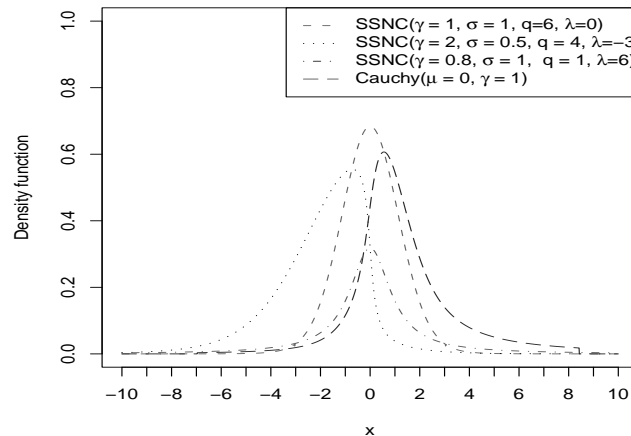


Figure 2 – Skew slash normal-Cauchy density functions for various values of parameters.

Remark 8. The skew slash normal-Cauchy distribution has heavier tails since the survival function decays at the rate of power function. The heavy tail characteristic makes these densities appropriate for modeling microarray, network delays, signals and noise, financial risk or interference which is impulsive in nature.

3. SKEW SLASH NORMAL-LAPLACE DISTRIBUTION

The skew slash normal-Laplace distribution can be defined as the distribution of the ratio $X = Y / U^{1/q}$ where Y is a skew normal-Laplace random variable and U is an independent uniform random variable over $(0, 1)$ and $q > 0$. It is denoted by $X \sim \text{SSNL}(\gamma, \sigma, \lambda; q)$.

Definition 3. A random variable X denoted by $X \sim SSNL(\gamma, \sigma, \lambda; q)$ is said to have a skew slash normal-Laplace distribution if its probability density function is

$$g(x; q) = \frac{q}{\sqrt{2\pi}\gamma} \int_0^1 v^q e^{-\frac{x^2 v^2}{2\gamma^2}} \begin{cases} \exp\left\{\frac{\lambda xv}{\sigma}\right\} dv, & \text{for } x \leq 0 \\ \left(2 - \exp\left(-\frac{\lambda xv}{\sigma}\right)\right) dv, & \text{for } x > 0 \end{cases} \quad (7)$$

The characteristic function of the skew slash normal-Laplace distribution is given by,

$$\varphi_X(t) = \int_0^1 \phi_Y(tu^{-1/q}) du$$

where $\phi_Y(\cdot)$ is the characteristic function of the skew normal-Laplace distribution (Nadarajah and Kotz (2003)) and is given by

$$\begin{aligned} \phi_Y(t) = E(e^{itY}) = & 2\exp\left(\frac{-\gamma^2 t^2}{2}\right) \Phi(i\gamma t) - \exp\left(\frac{\gamma^2 \left(\frac{\lambda}{\sigma} + it\right)^2}{2}\right) \Phi\left(-\gamma\left(\frac{\lambda}{\sigma} + it\right)\right) + \\ & \exp\left(\frac{\gamma^2 \left(\frac{\lambda}{\sigma} - it\right)^2}{2}\right) \Phi\left(-\gamma\left(\frac{\lambda}{\sigma} - it\right)\right) \end{aligned}$$

The mean of X is given by

$$E(X) = \frac{q}{q-1} \frac{2\lambda\gamma^2}{\sigma} \exp\left(\frac{\lambda^2\gamma^2}{2\sigma^2}\right) \Phi\left(-\frac{\lambda\gamma}{\sigma}\right), \quad q > 1$$

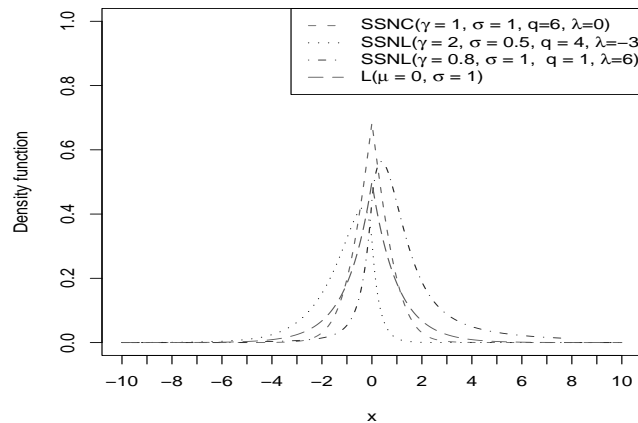


Figure 3 – Skew slash normal-Laplace density functions for various values of parameters.

Remark 9. Note that the skew slash normal-Laplace random variable in (7) is a scale mixture of the skew normal-Laplace random variable and so it can be represented as,

$$X | (U = u) \sim \text{SNL}(0, u^{-1/q} \gamma, u^{-1/q} \sigma, \lambda) \text{ with } U \sim U(0,1)$$

The probability density function (7) has the similar properties of the SSNN density function (5).

Figure 3 gives the probability density function of the skew slash normal-Laplace density for various values of the parameters.

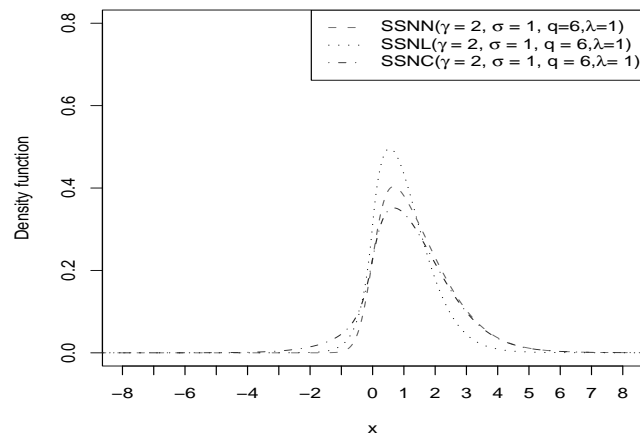


Figure 4 – SSNN density along with SSNL and SSNC.

From Figure 4, we can see that SSNL is more peaked than SSNN and SSNC. These distributions add skewness, heavier tails and peakedness in to the normal distribution. These characteristics make these densities appropriate for modeling microarray, network delays, signals and noise, financial risk or interference which is impulsive in nature.

Department of Statistics
St. Thomas College, Pala, Kerala, India

BINDU PUNATHUMPARAMBATH

ACKNOWLEDGMENT

The author is grateful to the Department of Science & Technology, Government of India, New Delhi, for financial support under the Women Scientist Scheme (WOS-A (2008)), Project No: SR/WOS-A/MS-09/2008.

REFERENCES

- A. AZZALINI (1985), *A class of distributions which includes the normal ones*, "Scandinavian Journal of Statistics", 2, pp. 171-178.
- A. K. GUPTA, F. C. CHANG, HUANG, W. J. (2002), *Some skew-symmetric models*, "Random Operators and Stochastic Equations", 10, pp. 133-140.
- K. KAFADAR (1982), *A biveight approach to the one-sample problem*, "Journal of American Statistical Association", 77, 416-424.
- K. KAFADAR (1988), *Slash Distribution*, "Encyclopedia of Statistical Sciences", 8.
- Y. MA, M.G. GENTON (2004), *A flexible class of skew-symmetric distributions*, "Scandinavian Journal of Statistics", 31, pp. 459-468.
- F. MOSTELLER, J.W. TUKEY (1977), *Data analysis and regression*, Addison-Wesley, Reading, MA.
- W.H. ROGERS, J.W. TUKEY (1972), *Understanding some long-tailed symmetrical distributions*, "Statistica Neerlandica", 26, pp. 211-226.
- S. NADARAJAH AND S. KOTZ (2003), *Skewed distributions generated by the normal kernel*, "Statistics and Probability letters", 65, pp. 269-277.
- J. WANG, M.G. GENTON (2006), *The multivariate skew-slash distribution*, "Journal of Statistical Planning and Inference", 136, pp. 209-220.

SUMMARY

A new family of skewed slash distributions generated by the normal kernel

The present paper is a generalization of the recent paper by Nadaraja and Kotz (2003) (*Skewed distributions generated by the normal kernel*, "Statistics & Probability Letters", 65, pp. 269-277). The new family of univariate skewed slash distributions generated by the normal kernel arises as the ratio of skewed distributions generated by the normal kernel and independent uniform power function distribution. The properties of the resulting distributions are studied. Normal, skew normal, slash (slash normal) and skew slash distributions are special cases of this new family. The normal distribution belongs to this family, since when the skewness parameter is zero and tail parameter tends to infinity the skew slash distributions generated by normal kernel reduces to the normal distribution. The slash normal family is also belongs to this family when the skewness parameter is zero. These distributions provide us alternative choices in simulation study and in particular, in fitting skewed data sets with heavy tails. We believe that the new class will be useful for analyzing data sets having skewness and heavy tails. Heavy-tailed distributions are commonly found in complex multi-component systems like ecological systems, microarray, biometry, economics, sociology, internet traffic, finance, business etc. We are working on maximum likelihood estimation of the parameters using EM algorithm and to apply our models for analysing the genetic data sets.