A Folded Normal Slash Distribution and Its Applications to Non-negative Measurements

Wenhao Gui[1]*, Pei-Hua Chen[2] and Haiyan Wu[3]
[1]*University of Minnesota Duluth*, [2]*National Chiao Tung University and*
[3]*Florida State University*

Abstract: We introduce a new class of the slash distribution using folded normal distribution. The proposed model defined on non-negative measurements extends the slashed half normal distribution and has higher kurtosis than the ordinary half normal distribution. We study the characterization and properties involving moments and some measures based on moments of this distribution. Finally, we illustrate the proposed model with a simulation study and a real application.

Key words: Folded normal distribution, kurtosis, maximum likelihood, skewness, slash distribution.

1. Introduction

The folded normal distribution, proposed by Leone *et al.* (1961), is often used when the measurement system produces only non-negative measurements, from an otherwise normally distributed process. Given a normally distributed random variable $W \sim \mathcal{N}(\mu, \sigma)$ with mean μ and standard deviation σ, its absolute value $X = |W|$ has a folded normal distribution. The density function of the folded normal distribution is given by, for $x \geq 0$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}[e^{-\frac{(x+\mu)^2}{2\sigma^2}} + e^{-\frac{(x-\mu)^2}{2\sigma^2}}]. \qquad (1)$$

Replacing the parameters (μ, σ) by (θ, σ), where $\theta = \mu/\sigma$, the density can be expressed as another version, denoted as $\mathcal{FN}(\theta, \sigma)$, see Johnson (1962),

$$f(x) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}}e^{-\frac{\theta^2}{2}}e^{-\frac{x^2}{2\sigma^2}}\cosh(\frac{\theta x}{\sigma}), \qquad x \geq 0. \qquad (2)$$

---

*Corresponding author.

Such a case may be encountered if measurements are frequently recorded without their algebraic sign. When $\mu = 0$ ($\theta = 0$), the folded normal random variable reduces to the well known half normal random variable. The folded normal $|W|$ and truncated normal $W|W > 0$ are equivalent in this case. For $\mu \neq 0$, the folded normal and truncated normal are totally different since the distribution of W is not symmetric at zero.

Leone *et al.* (1961) described methods for estimating the mean and standard deviation of the normal distribution based on estimates of the mean and standard deviation determined from the folded normal distribution. Elandt (1961) derived the general formula for the moments of the folded normal distribution. Johnson (1962) studied the accuracy of estimation of the parameters by maximum likelihood. Lin (2005) proposed a generalized folded normal distribution and applied it to the process capability measures.

On the other hand, a random variable S has a standard slash distribution $\mathcal{SL}(q)$ with parameter $q > 0$ if S can be represented as

$$S = \frac{Z}{U^{\frac{1}{q}}}, \tag{3}$$

where $Z \sim \mathcal{N}(0,1)$ and $U \sim \mathcal{U}(0,1)$ are independent. This distribution was named by Rogers and Tukey (1972). It generalizes normality and has been much studied in the statistical literature.

The $\mathcal{SL}(q)$ density function is given by

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

where $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ denotes the standard normal density function.

For the limit case $q \to \infty$, $\mathcal{SL}(q)$ yields the standard normal distribution Z. Let $q = 1$, the canonical slash distribution follows and it has the following density, see Johnson *et al.* (1995),

$$f(x) = \begin{cases} \frac{\phi(0)-\phi(x)}{x^2}, & x \neq 0, \\ \frac{1}{2}\phi(0), & x = 0. \end{cases} \tag{5}$$

It is known that the standard slash density has heavier tails and larger kurtosis than those of the normal distribution. It has been very popular in robust statistical analysis and actively studied by some authors.

Rogers and Tukey (1972) and Mosteller and Tukey (1977) investigated the general properties of this canonical slash distribution. Kafadar (1982) studied the maximum likelihood estimates of the location and scale parameters.

Gómez *et al.* (2007a) replaced standard normal random variable Z by an elliptical distribution and defined a new family of slash distributions. They studied its general properties of the resulting families, including their moments. Genc (2007) derived the univariate slash by a scale mixtured exponential power distribution and investigated asymptotically the bias properties of the estimators. Wang and Genton (2006) proposed the multivariate skew version of this distribution and examined its properties and inferences. They substituted the standard normal random variable Z by a skew normal distribution introduced by Azzalini (1985) to define a skew extension of the slash distribution. Olmos *et al.* (2011) introduced the slashed half normal distribution by a scale mixtured half normal distribution and showed the resulting distribution has higher kurtosis than the ordinary half normal distribution.

Since the folded normal distribution generalizes the half normal distribution, it is naturally to define a slash distribution based on it in case there is great diversity in the distribution's tail. In our work, we introduce a new slash distribution, so-called the folded normal slash (FNS) distribution for the first time, which contains as sub-models the half normal (HN) and the folded normal (FN) distributions etc.

The rest of this paper is organized as follows: in Section 2, we propose the new slash distribution and investigate its properties, including the stochastic representation etc. Section 3 discusses the inference, moments and maximum likelihood estimation for the parameters. Simulation studies are performed in Section 4. Section 5 gives a real illustrative application and reports the results. Section 6 concludes our work.

2. Folded Normal Slash Distribution

2.1 Stochastic Representation

Definition 2.1. A random variable Y has a folded normal slash distribution if it can be represented as the ratio

$$Y = \frac{X}{U^{\frac{1}{q}}}, \tag{6}$$

where $X \sim \mathcal{FN}(\theta, \sigma)$ defined in (2) and $U \sim \mathcal{U}(0, 1)$ are independent, $-\infty < \theta < \infty$, $\sigma > 0$, $q > 0$. We denote it as $Y \sim \mathcal{FNS}(\theta, \sigma, q)$.

Proposition 2.2. Let $Y \sim \mathcal{FNS}(\theta, \sigma, q)$. Then, the density function of Y is given by

$$f_Y(y) = \frac{q\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{\theta^2}{2}} \int_0^1 e^{-\frac{y^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y t}{\sigma}) t^q dt \quad \text{for} \quad y \geq 0, \tag{7}$$

where $-\infty < \theta < \infty$, $\sigma > 0$, $q > 0$.

Particularly, if $\theta = 0$, the density function (7) reduces to

$$f_Y(y) = \frac{q\sqrt{2}}{\sigma\sqrt{\pi}} \int_0^1 e^{-\frac{y^2 t^2}{2\sigma^2}} t^q dt,$$

which is the density function for the slashed half normal distribution discussed by Olmos *et al.* (2011). As $q \to \infty$,

$$\lim_{q\to\infty} f_Y(y) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{\theta^2}{2}} e^{-\frac{y^2}{2\sigma^2}} \cosh(\frac{\theta y}{\sigma}).$$

The limit case of the folded normal slash distribution is the folded normal distribution. For $\sigma = q = 1$, the canonical case is obtained.

Figure 1 shows some density functions of the canonical cases of folded normal slash distribution with various parameters.
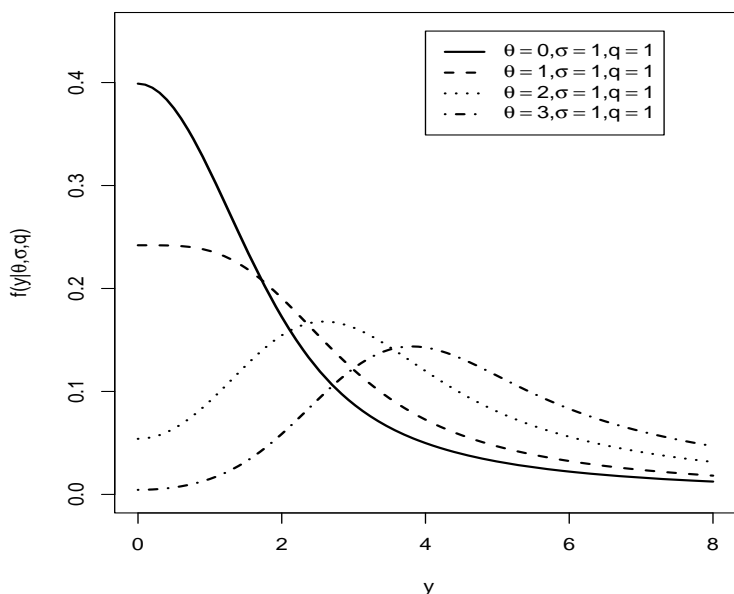

Figure 1: The density function of $\mathcal{FNS}(\theta, \sigma, q)$ with various parameters

The cumulative distribution function of the folded normal slash distribution $Y \sim \mathcal{FNS}(\theta, \sigma, q)$ is given as follows. For $y \geq 0$,

$$F_Y(y) = \int_0^y f_Y(u) du = \frac{q}{2} \int_0^1 t^{q-1} \left[\mathrm{erf}(\frac{yt + \theta\sigma}{\sqrt{2}\sigma}) + \mathrm{erf}(\frac{yt - \theta\sigma}{\sqrt{2}\sigma}) \right] dt, \quad (8)$$

where erf() is the Gauss error function, defined as $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

Proposition 2.3. Let $Y|U = u \sim \mathcal{FN}(\theta, \sigma u^{-1/q})$ and $U \sim \mathcal{U}(0,1)$, then $Y \sim \mathcal{FNS}(\theta, \sigma, q)$.

Proposition 2.3 shows that the folded normal slash distribution can be represented as a scale mixture of a folded normal distribution and uniform distribution. The representation provides another way besides the definition (6) to generate random numbers from the folded normal slash distribution $\mathcal{FNS}(\theta, \sigma, q)$, which is very useful to conduct simulations. It can also be used to study farther properties of the folded normal slash distribution.

2.2 Moments and Measures Based on Moments

Elandt (1961) derived the moments of folded normal distribution. The following notations are used.

$$I_j(a) = \frac{1}{\sqrt{2\pi}} \int_a^\infty y^j e^{-\frac{y^2}{2}} dy \quad \text{for} \quad j = 0, 1, \cdots . \tag{9}$$

Proposition 2.4. Let $X \sim \mathcal{FN}(\theta, \sigma)$, then the kth non-central moments are given by

$$E(X^k) = \sigma^k \sum_{j=0}^{k} \binom{k}{j} \theta^{k-j} [I_j(-\theta) + (-1)^{k-j} I_j(\theta)], \tag{10}$$

for $k = 1, 2, \cdots$.

Proof. See Elandt (1961).   □

Proposition 2.5. Let $Y \sim \mathcal{FNS}(\theta, \sigma, q)$. For $k = 1, 2, \cdots$ and $q > k$, the kth non-central moment of Y is given by

$$\mu_k = E(Y^k) = \frac{q\sigma^k}{q-k} \sum_{j=0}^{k} \binom{k}{j} \theta^{k-j} [I_j(-\theta) + (-1)^{k-j} I_j(\theta)]. \tag{11}$$

Proof. From the stochastic representation defined in (6) and the results in (10), the claim follows in a straightforward manner.

$$\mu_k = E(Y^k) = E[(XU^{-\frac{1}{q}})^k] = E(X^k) \times E(U^{-\frac{k}{q}}) = \frac{q}{q-k} E(X^k).$$

□

The following results are immediate consequences of (11).

Corollary 2.6. Let $Y \sim \mathcal{FNS}(\theta, \sigma, q)$. The mean and variance of Y are given by

$$E(Y) = \frac{q\sigma}{q-1}[2\phi(\theta) + 2\theta\Phi(\theta) - \theta], \quad q > 1 \quad \text{and} \tag{12}$$

$$\text{Var}(Y) = \frac{q(\theta^2 + 1)\sigma^2}{q-2} - \frac{q^2\sigma^2}{(q-1)^2}[2\phi(\theta) + 2\theta\Phi(\theta) - \theta]^2, \quad q > 2, \tag{13}$$

where ϕ and Φ are the probability density and cumulative distribution functions of standard normal random variable. For the standardized skewness and kurtosis coefficients,

$$\sqrt{\beta_1} = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{3/2}},$$

and

$$\beta_2 = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1^2)^2},$$

we have the following results.

Corollary 2.7. Let $Y \sim \mathcal{FNS}(\theta, \sigma, q)$. The skewness and kurtosis coefficients of Y are given by

$$\sqrt{\beta_1} = \frac{\left\{ \begin{array}{c} \frac{q[-\theta^3 + 2\theta^3\Phi(\theta) - 3\theta + 6\theta\Phi(\theta) + 4\phi(\theta) + 2\theta^2\phi(\theta)]}{q-3} - \frac{3q^2[2\phi(\theta) + 2\theta\Phi(\theta) - \theta](\theta^2 + 1)}{(q-1)(q-2)} \\ + \frac{2q^3[2\phi(\theta) + 2\Phi(\theta)\theta - \theta]^3}{(q-1)^3} \end{array} \right\}}{\left[\frac{q(\theta^2 + 1)}{q-2} - \frac{q^2}{(q-1)^2}[2\phi(\theta) + 2\theta\Phi(\theta) - \theta]^2 \right]^{3/2}}, \tag{14}$$

for $q > 3$;

$$\beta_2 = \frac{\left\{ \begin{array}{c} \frac{q(\theta^4 + 6\theta^2 + 3)}{q-4} - \frac{4q^2[2\phi(\theta) + 2\theta\Phi(\theta) - \theta][-\theta^3 + 2\theta^3\Phi(\theta) - 3\theta + 6\theta\Phi(\theta) + 4\phi(\theta) + 2\theta^2\phi(\theta)]}{(q-1)(q-3)} \\ + \frac{6q^3[2\phi(\theta) + 2\theta\Phi(\theta) - \theta]^2(\theta^2 + 1)}{(q-1)^2(q-2)} - \frac{3q^4[2\phi(\theta) + 2\theta\Phi(\theta) - \theta]^4}{(q-1)^4} \end{array} \right\}}{\left[\frac{q(\theta^2 + 1)}{q-2} - \frac{q^2}{(q-1)^2}[2\phi(\theta) + 2\theta\Phi(\theta) - \theta]^2 \right]^2}, \tag{15}$$

for $q > 4$.

As $q \to \infty$, the skewness and kurtosis coefficients converge.

$$\sqrt{\beta_1} \to \frac{\left\{ \begin{array}{c} [-\theta^3 + 2\theta^3\Phi(\theta) - 3\theta + 6\theta\Phi(\theta) + 4\phi(\theta) + 2\theta^2\phi(\theta)] - 3[2\phi(\theta) + 2\theta\Phi(\theta) - \theta](\theta^2 + 1) \\ + 2[2\phi(\theta) + 2\Phi(\theta)\theta - \theta]^3 \end{array} \right\}}{[(\theta^2 + 1) - [2\phi(\theta) + 2\theta\Phi(\theta) - \theta]^2]^{3/2}},$$

and

$$\beta_2 \to \frac{\left\{\begin{array}{c}(\theta^4+6\theta^2+3)-4[2\phi(\theta)+2\theta\Phi(\theta)-\theta][-\theta^3+2\theta^3\Phi(\theta)-3\theta+6\theta\Phi(\theta)+4\phi(\theta)+2\theta^2\phi(\theta)]\\6[2\phi(\theta)+2\theta\Phi(\theta)-\theta]^2(\theta^2+1)-3[2\phi(\theta)+2\theta\Phi(\theta)-\theta]^4\end{array}\right\}}{[(\theta^2+1)-[2\phi(\theta)+2\theta\Phi(\theta)-\theta]^2]^2},$$

which are the skewness and kurtosis coefficients for the folded normal distribution $\mathcal{FN}(\theta,\sigma)$.

Figure 2 shows the skewness and kurtosis coefficients with various parameters for the $\mathcal{FNS}(\theta,\sigma,q)$ model. The skewness and kurtosis coefficients decrease as q increases. Also, for q is large, the skewness and kurtosis coefficients decrease as θ increases. The parameter σ does not affect the two coefficients.

(a) Skewness coefficient   (b) Kurtosis coefficient
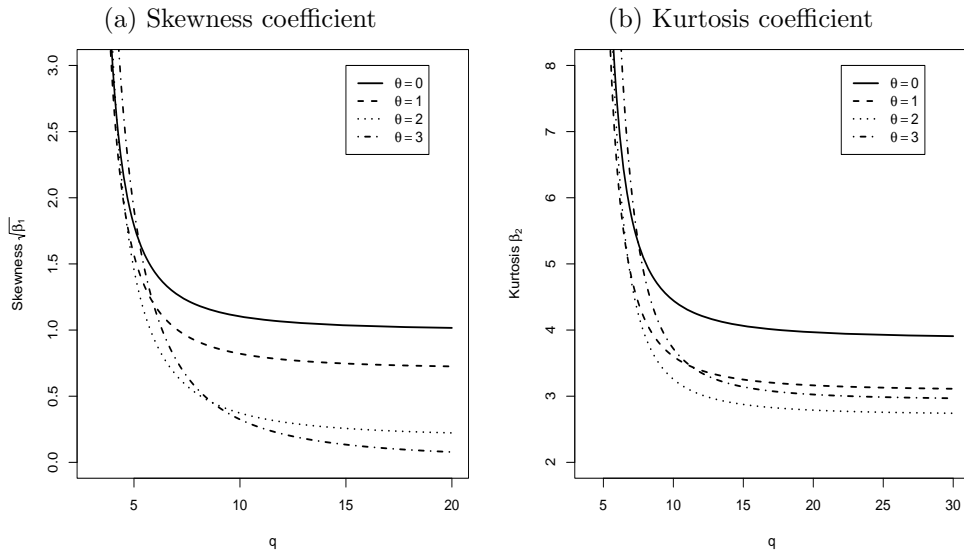

Figure 2: The plot for the skewness $\sqrt{\beta_1}$ and kurtosis coefficient β_2 with various parameters

3. Maximum Likelihood (ML) Inference

In this section, we consider the maximum likelihood estimation about the parameters (θ,σ,q) of the FNS model. Suppose $y_1, y_2, \cdots, y_n$ is a random sample of size n from the folded normal slash distribution $\mathcal{FNS}(\theta,\sigma,q)$. Then the log-likelihood function is given by

$$l = \log \prod_{i=1}^{n} f_Y(y_i)$$

$$= n\log q - n\log \sigma + \frac{n}{2}\log(\frac{2}{\pi}) - \frac{n\theta^2}{2} + \sum_{i=1}^{n} \log \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt.$$

(16)

The estimates of the parameters maximize the likelihood function. Taking the partial derivatives of the log-likelihood function with respect to θ, σ, q respectively and equalizing the obtained expressions to zero yield to likelihood equations. However, they cannot lead to explicit analytical solutions for model parameters. Thus, the estimates can be obtained by means of numerical procedures such as Newton-Raphson method. The program R provides the nonlinear optimization routine optim for solving such problems.

It is known that as the sample size increases, the distribution of the MLE tends to the normal distribution with mean (θ, σ, q) and covariance matrix equal to the inverse of the Fisher information matrix. The score vector and Hessian matrix are given in Appendix A.2. However, the log-likelihood function given in (16) is a complex expression. It is not generally possible to derive the Fisher information matrix. Thus, the theoretical properties (asymptotically normal, unbiased etc) of the MLE estimators are not easily obtained. We study the properties of the estimators numerically.

4. Simulation Study

In this section, we perform a simulation to illustrate the behavior of the MLE estimators for parameters θ, σ and q. We first generate 500 samples of size $n = 50$ and $n = 100$ from the $\mathcal{FNS}(\theta, \sigma, q)$ distribution for fixed parameters. The random numbers can be generated as follows:

Simulate $W \sim \mathcal{N}(\theta\sigma, \sigma)$ and $U \sim \mathcal{U}(0, 1)$, compute $Y = |W|U^{-1/q}$, it follows that $Y \sim \mathcal{FNS}(\theta, \sigma, q)$.

The estimates are computed by the optim function which uses L-BFGS-B method in software R. The empirical means and standard deviations of the estimates are presented in Table 1. From Table 1, the parameters are well estimated and the estimates are asymptotically unbiased.

In addition, a simulation is presented illustrating the sample properties of the moments. We consider the skewness $\sqrt{\beta_1}$ and kurtosis β_2 coefficients derived in Section 2. We generate 500 samples of size $n = 1000$ and $n = 2000$ from the $\mathcal{FNS}(\theta, \sigma, q)$ distribution for fixed parameters. For each generated sample, $y_1, \cdots, y_n$, the sample skewness $\sqrt{b_1}$ and sample kurtosis b_2 are computed by

$$\sqrt{b_1} = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^3}{(\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2)^{3/2}},$$

and

$$b_2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^4}{(\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2)^2}.$$

The mean values of the sample coefficients and simulated standard deviation (SD) are obtained and reported in Table 2. From Table 2, we can see that the sample

Table 1: Empirical means and SD for the MLE estimators of θ, σ and q

			$n = 50$			$n = 100$		
θ	σ	q	$\hat{\theta}$ (SD)	$\hat{\sigma}$ (SD)	$\hat{q}$ (SD)	$\hat{\theta}$ (SD)	$\hat{\sigma}$ (SD)	$\hat{q}$ (SD)
2	2	2	2.0663 (0.4195)	2.0607 (0.6097)	2.2781 (0.8620)	2.0171 (0.2763)	2.0454 (0.3764)	2.1087 (0.4324)
2	2	4	2.0699 (0.3778)	1.9969 (0.4499)	4.5470 (1.9785)	1.9917 (0.2353)	2.0488 (0.2973)	4.4275 (1.4005)
2	4	2	2.0530 (0.4666)	4.1048 (1.1135)	2.3916 (1.2174)	2.0080 (0.2917)	4.1027 (0.7721)	2.1151 (0.4580)
2	4	4	2.0411 (0.3638)	4.0989 (0.8699)	4.7641 (1.9402)	2.0023 (0.2480)	4.0721 (0.5756)	4.4270 (1.4590)
4	2	2	4.2147 (0.9060)	2.0293 (0.5031)	2.2102 (0.5611)	4.1144 (0.5633)	2.0118 (0.3303)	2.0945 (0.3212)
4	2	3	4.2518 (0.8607)	1.9789 (0.4714)	3.2708 (1.0197)	4.0928 (0.5379)	1.9996 (0.3107)	3.1376 (0.5808)
4	3	2	4.2697 (0.9880)	2.9658 (0.7052)	2.1169 (0.4764)	4.1118 (0.5938)	3.0041 (0.5101)	2.0683 (0.3155)
4	3	3	4.2206 (0.8920)	3.0076 (0.7136)	3.3026 (1.0749)	4.1519 (0.5767)	2.9546 (0.4666)	3.1078 (0.5856)

Table 2: Empirical means and SD for the skewness and kurtosis coefficients

				$n = 1000$		$n = 2000$	
θ	q	$\sqrt{\beta_1}$	β_2	$\sqrt{b_1}$ (SD)	b_2 (SD)	$\sqrt{b_1}$ (SD)	b_2 (SD)
1	5	1.5660	12.3408	1.5419 (0.4479)	10.5885 (0.6326)	1.5626 (0.5009)	11.5151 (0.5895)
2	5	1.4569	14.9363	1.4120 (0.5872)	12.1546 (0.8038)	1.4517 (0.5120)	13.7815 (0.6401)
3	5	1.9121	22.6475	1.8661 (0.9404)	18.5366 (0.7943)	1.8765 (0.7646)	19.2365 (0.6059)
1	6	1.1528	9.5588	1.1380 (0.9123)	8.9319 (0.9798)	1.1473 (0.3547)	9.2822 (0.7556)
2	6	0.9134	6.8155	0.9062 (0.2117)	6.5632 (0.9473)	0.9104 (0.1675)	6.6623 (0.4999)
3	6	1.1528	9.5588	1.1440 (0.3154)	9.1652 (0.8501)	1.1504 (0.2598)	9.4080 (0.6854)

skewness and kurtosis coefficients converge to the population coefficients as expected.

In the following, we conduct a simulation to compare the behavior of the moment estimator (MME) and the maximum likelihood estimator (MLE). We generate 500 samples of size $n = 100$ from the $\mathcal{FNS}(\theta, \sigma, q)$ distribution for fixed parameters. For each generated sample, we compute the ML estimates as well as the sample skewness and kurtosis coefficients. The results are displayed in Table 3. It can be seen that the moment estimators are less efficient than the MLE's. For small sample size such as $n = 100$, the MLE performs well and the estimates are very close to the true parameters. However, for the measures based on the MME, their values deviate from their population values and their behavior is poor. In fact, sample size has a big impact on both skewness and kurtosis.

Table 3: A simulation comparison between MLE and MME ($n = 100$)

θ	σ	q	$\hat{\theta}$ (SD)	$\hat{\sigma}$ (SD)	$\hat{q}$ (SD)	$\sqrt{\beta_1}$	β_2	$\sqrt{b_1}$ (SD)	b_2 (SD)
2	1	5	2.0048 (0.2494)	1.0355 (0.1742)	5.1891 (0.5807)	1.4569	14.9363	1.0213 (0.8435)	5.9111 (2.0302)
2	1	6	2.0226 (0.2407)	1.0328 (0.1503)	5.7785 (0.5163)	0.9134	6.8155	0.6463 (0.4736)	4.0805 (2.2514)
2	1	7	2.0259 (0.2115)	1.0243 (0.1386)	6.7216 (0.5688)	0.6574	4.7334	0.5430 (0.5035)	3.8703 (1.7476)
3	1	5	3.0539 (0.3138)	1.0045 (0.1391)	5.3766 (0.8109)	1.9121	22.6475	1.2076 (0.9245)	12.1143 (3.6377)
3	1	6	3.0216 (0.2999)	1.0212 (0.1301)	6.2134 (0.5374)	1.1528	9.5588	0.7710 (0.7279)	5.1320 (1.9533)
3	1	7	3.0366 (0.3201)	1.0137 (0.1271)	6.7561 (0.5276)	0.7724	6.1320	0.5738 (0.5435)	4.2818 (1.1654)

5. Real Data Illustration

In this section, we consider the Australian Sports data. The data are the plasma ferritin concentration measurements of 202 athletes collected at the Australian Institute of Sport. This data set has been well studied by several authors, see Azzalini and Dalla Valle (1996), Cook and Weisberc (1994) and Elal-Olivero et al. (2009).

Table 4 summarizes descriptive statistics of the data set. This data set indicates non negative asymmetry. We fit the data set with the half normal(HN), the folded normal(FN) and the folded normal slash(FNS) distributions, respectively, using maximum likelihood method. The results are shown in Table 5. The usual Akaike information criterion (AIC) introduced by Akaike (1973) and Bayesian information criterion (BIC) proposed by Schwarz (1978) to measure of the goodness of fit are also computed. $\text{AIC} = 2k - 2\log L$ and $\text{BIC} = k\log n - 2\log L$.

where k is the number of parameters in the distribution and L is the maximized value of the likelihood function. The results show that FNS model fits best. Figure (3a) and (3b) display the fitted models using the MLE estimates.

Table 4: Summary for the plasma ferritin concentration measurements

sample size	mean	standard deviation	$\sqrt{b_1}$	b_2
202	76.88	47.50	1.28	4.42

Table 5: Maximum likelihood parameter estimates (with (SD)) of the HN, FN and FNS models for the plasma ferritin concentration data

Model	$\hat{\theta}$	$\hat{\sigma}$	$\hat{q}$	loglik	AIC	BIC
HN	–	76.9436	–	-1062.037	2126.074	2129.382
		(3.0588)				
FN	1.7288	41.4799	–	-1065.875	2135.75	2142.367
	(0.1114)	(1.7256)				
FNS	2.4227	19.1816	2.3775	-1037.344	2080.688	2090.613
	(0.2387)	(2.8424)	(0.3778)			

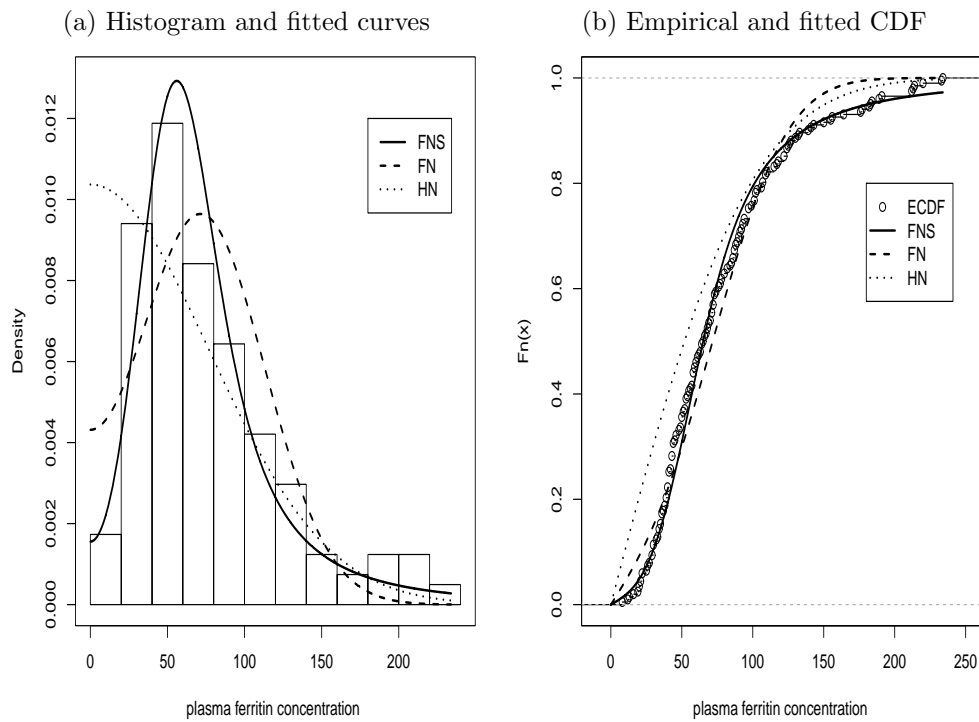(a) Histogram and fitted curves   (b) Empirical and fitted CDF

Figure 3: Models fitted for the plasma ferritin concentration data set

6. Concluding Remarks

In this article, we have studied the folded normal slash distribution, denoted by $\mathcal{FNS}(\theta, \sigma, q)$. It is defined to be the quotient of two independent random variables, a folded normal random variable and a power of the uniform distribution. This proposed non-negative distribution generalizes the folded normal, the half normal distribution etc.

Probabilistic and inferential properties are studied. A simulation is conducted and demonstrates the good performance of the maximum likelihood estimators. We apply the model to a real dataset and the results demonstrate that the proposed model is appropriate and flexible for non-negative data.

There are a number of possible extensions of the current work. Lin *et al.* (2007b) addressed the problem of analyzing a mixture of skew normal distributions from the likelihood-based and Bayesian perspectives, respectively. Lin *et al.* (2007a) proposed a robust mixture framework based on the skew t distribution to efficiently deal with heavy-tailedness, extra skewness and multimodality in a wide range of settings. Ho *et al.* (2012) dealt with the problem of maximum likelihood estimation for a mixture of skew Student-t-normal distributions. Thus, mixture modeling using the proposed distributions is the most natural extension. Other extensions of the current work include, for example, a generalization of the folded normal slash distribution to multivariate settings.

Appendix

A.1 Proofs of Propositions

(1) Proof of Proposition 2.2.

Proof. From (2), the joint probability density function of X and U is given by

$$g(x, u) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{\theta^2}{2}} e^{-\frac{x^2}{2\sigma^2}} \cosh(\frac{\theta x}{\sigma}) \ \ \text{for} \ \ x \geq 0, \ 0 < u < 1.$$

Using the transformation: $y = \frac{x}{u^{1/q}}, u = u$, the joint probability density function of Y and U is given by

$$h(y, u) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{\theta^2}{2}} e^{-\frac{y^2 u^{2/q}}{2\sigma^2}} \cosh(\frac{\theta y u^{1/q}}{\sigma}) u^{\frac{1}{q}} \ \ \text{for} \ \ y \geq 0, \ 0 < u < 1.$$

The marginal density function of Y is given by

$$f_Y(y) = \int_0^1 \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{\theta^2}{2}} e^{-\frac{y^2 u^{2/q}}{2\sigma^2}} \cosh(\frac{\theta y u^{1/q}}{\sigma}) u^{\frac{1}{q}} du.$$

After changing the variable into $t = u^{\frac{1}{q}}$, the density function will be obtained as stated. $\qquad\square$

(2) Proof of Proposition 2.3.

Proof.

$$f_Y(y) = \int_0^1 f_{Y|U}(y|u) f_U(u) du = \int_0^1 \frac{\sqrt{2}}{\sigma u^{-1/q}\sqrt{\pi}} e^{-\frac{\theta^2}{2}} e^{-\frac{y^2}{2(\sigma u^{-1/q})^2}} \cosh(\frac{\theta y}{\sigma u^{-1/q}}) du.$$

$\qquad\square$

A.2 Score Vector and Hessian Matrix

Suppose $y_1, y_2, \cdots, y_n$ is a random sample drawn from the folded normal slash distribution $\mathcal{FNS}(\theta, \sigma, q)$, then the log-likelihood function is given by (16). The elements of the score vector are obtained by differentiation

$$l_\theta = -n\theta + \sum_{i=1}^n \frac{y_i \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \sinh(\frac{\theta y_i t}{\sigma}) t^{q+1} dt}{\sigma \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt},$$

$$l_\sigma = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} [y_i^2 t^2 \cosh(\frac{\theta y_i t}{\sigma}) - \sigma \theta y_i t \sinh(\frac{\theta y_i t}{\sigma})] t^q dt}{\sigma^3 \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt},$$

$$l_q = \frac{n}{q} + \sum_{i=1}^n \frac{\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q \log t dt}{\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt}.$$

The Hessian matrix, second partial derivatives of the log-likelihood, is given by

$$H = \begin{pmatrix} l_{\theta\theta} & l_{\theta\sigma} & l_{\theta q} \\ l_{\sigma\theta} & l_{\sigma\sigma} & l_{\sigma q} \\ l_{q\theta} & l_{q\sigma} & l_{qq} \end{pmatrix},$$

where

$$l_{\theta\theta} = -n + \sum_{i=1}^n \frac{y_i^2 \left\{ \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^{q+2} dt \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt \atop -(\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \sinh(\frac{\theta y_i t}{\sigma}) t^{q+1} dt)^2 \right\}}{\sigma^2 (\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt)^2},$$

$$l_{\theta\sigma} = l_{\sigma\theta}$$

$$= \sum_{i=1}^{n} \frac{\left\{\begin{array}{c} \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} [\sigma^{-1} y_i^3 t^3 \sinh(\frac{\theta y_i t}{\sigma}) - \sigma y_i t \sinh(\frac{\theta y_i t}{\sigma}) - \theta y_i^2 t^2 \cosh(\frac{\theta y_i t}{\sigma})] t^q dt \times \\ \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt \\ -\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} [y_i^2 t^2 \cosh(\frac{\theta y_i t}{\sigma}) - \sigma\theta y_i t \sinh(\frac{\theta y_i t}{\sigma})] t^q dt \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \frac{y_i t}{\sigma} \sinh(\frac{\theta y_i t}{\sigma}) t^q dt \end{array}\right\}}{\sigma^3 [\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt]^2},$$

$$l_{\theta q} = l_{q\theta} = \sum_{i=1}^{n} \frac{\left\{\begin{array}{c} \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} y_i \sinh(\frac{\theta y_i t}{\sigma}) t^{q+1} \log t dt \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt \\ -\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q \log t dt \int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} y_i \sinh(\frac{\theta y_i t}{\sigma}) t^{q+1} dt \end{array}\right\}}{\sigma [\int_0^1 e^{-\frac{y_i^2 t^2}{2\sigma^2}} \cosh(\frac{\theta y_i t}{\sigma}) t^q dt]^2},$$

$$l_{\sigma q} = l_{q\sigma}$$

$$= -\sum_{i=1}^{n} \frac{\int_0^1 e^{-\frac{t^2 y^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) \log t \, dt \int_0^1 [\frac{e^{-\frac{t^2 y^2}{2\sigma^2}} t^{2+q} y^2 \cosh(\frac{t\theta y_i}{\sigma})}{\sigma^3} - \frac{e^{-\frac{t^2 y^2}{2\sigma^2}} t^{1+q} \theta y \sinh(\frac{t\theta y_i}{\sigma})}{\sigma^2}] \, dt}{\left(\int_0^1 e^{-\frac{t^2 y^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) \, dt\right)^2}$$

$$+ \sum_{i=1}^{n} \frac{\int_0^1 [\frac{e^{-\frac{t^2 y^2}{2\sigma^2}} t^{2+q} y^2 \cosh(\frac{t\theta y_i}{\sigma}) \log t}{\sigma^3} - \frac{e^{-\frac{t^2 y^2}{2\sigma^2}} t^{1+q} \theta y \log t \sinh(\frac{t\theta y_i}{\sigma})}{\sigma^2}] \, dt}{\int_0^1 e^{-\frac{t^2 y^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) \, dt},$$

$$l_{qq} = -\frac{n}{q^2} - \sum_{i=1}^{n} \left\{ \frac{[\int_0^1 e^{-\frac{t^2 y_i^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) \log t \, dt]^2}{[\int_0^1 e^{-\frac{t^2 y_i^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) \, dt]^2} + \frac{\int_0^1 e^{-\frac{t^2 y_i^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) [\log t]^2 \, dt}{\int_0^1 e^{-\frac{t^2 y_i^2}{2\sigma^2}} t^q \cosh(\frac{t\theta y_i}{\sigma}) \, dt} \right\}.$$

Acknowledgements

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (Edited by B. N. Petrov and F. Csaki), 267-281. Akademiai Kiado, Budapest.

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

Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715-726.

Box, G. and Tiao, G. (1962). A further look at robustness via Bayes's theorem. *Biometrika* **49**, 419-432.

Cook, R. D. and Weisberc, S. (1994). *An Introduction to Regression Graphics.* Wiley, New York.

Elal-Olivero, D., Olivares-Pacheco, J., Gómez, H. W. and Bolfarine, H. (2009). A new class of non negative distributions generated by symmetric distributions. *Communications in Statistics - Theory and Methods* **38**, 993-1008.

Elandt, R. C. (1961). The folded normal distribution: two methods of estimating parameters from moments. *Technometrics* **3**, 551-562.

Genc, A. (2007). A generalization of the univariate slash by a scale-mixtured exponential power distribution. *Communications in Statistics - Simulation and Computation* **36**, 937-947.

Gómez, H. W., Quintana, F. A. and Torres, F. J. (2007a). A new family of slash-distributions with elliptical contours. *Statistics & Probability Letters* **77**, 717-725.

Gómez, H. W., Venegas, O. and Bolfarine, H. (2007b). Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics* **18**, 395-407.

Goodman, I. R. and Kotz, S. (1973). Multivariate θ-generalized normal distributions. *Journal of Multivariate Analysis* **3**, 204-219.

Ho, H. J., Pyne, S. and Lin, T. I. (2012). Maximum likelihood inference for mixtures of skew student-t-normal distributions through practical EM-type algorithms. *Statistics and Computing* **22**, 287-299.

Jamshidian, M. (2001). A note on parameter and standard error estimation in adaptive robust regression. *Journal of Statistical Computation and Simulation* **71**, 11-27.

Johnson, N. L. (1962). The folded normal distribution: accuracy of estimation by maximum likelihood. *Technometrics* **4**, 249-256.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions.* Wiley, New York.

Kafadar, K. (1982). A biweight approach to the one-sample problem. *Journal of the American Statistical Association* **77**, 416-424.

Kashid, D. N. and Kulkarni, S. R. (2003). Subset selection in multiple linear regression with heavy tailed error distribution. *Journal of Statistical Computation and Simulation* **73**, 791-805.

Leone, F. C., Nelson, L. S. and Nottingham, R. B. (1961). The folded normal distribution. *Technometrics* **3**, 543-550.

Lin, P. C. (2005). Application of the generalized folded-normal distribution to the process capability measures. *The International Journal of Advanced Manufacturing Technology* **26**, 825-830.

Lin, T. I., Lee, J. C. and Hsieh, W. J. (2007a). Robust mixture modeling using the skew *t* distribution. *Statistics and Computing* **17**, 81-92.

Lin, T. I., Lee, J. and Yen, S. Y. (2007b). Finite mixture modelling using the skew normal distribution. *Statistica Sinica* **17**, 909-927.

Ma, Y. and Genton, M. G. (2004). Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics* **31**, 459-468.

Morgenthaler, S. (1986). Robust confidence intervals for a location parameter: the configural approach. *Journal of the American Statistical Association* **81**, 518-525.

Mosteller, F. and Tukey, J. W. (1977). *Data Analysis and Regression: A Second Course in Statistics.* Addison-Wesley Series in Behavioral Science: Quantitative Methods. Reading, Addison-Wesley, Massachusetts.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics* **32**, 685-694.

Olmos, N. M., Varela, H., Gómez, H. W. and Bolfarine, H. (2012). An extension of the half-normal distribution. *Statistical Papers* **53**, 875-886.

Pescim, R. R., Demétrio, C. G. B., Cordeiro, G. M., Ortega, E. M. M. and Urbano, M. R. (2010). The beta generalized half-normal distribution. *Computational Statistics & Data Analysis* **54**, 945-957.

Pewsey, A. (2002). Large-sample inference for the general half-normal distribution. *Communications in Statistics - Theory and Methods* **31**, 1045-1054.

Pewsey, A. (2004). Improved likelihood based inference for the general half-normal distribution. *Communications in Statistics - Theory and Methods* **33**, 197-204.

Rogers, W. H. and Tukey, J. W. (1972). Understanding some long-tailed symmetrical distributions. *Statistica Neerlandica* **26**, 211-226.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461-464.

Wang, J. and Genton, M. G. (2006). The multivariate skew-slash distribution. *Journal of Statistical Planning and Inference* **136**, 209-220.

Wenhao Gui
Department of Mathematics and Statistics
University of Minnesota Duluth
Duluth, MN 55812, USA
wgui@d.umn.edu

Pei-Hua Chen
Department of Management Science
National Chiao Tung University
1001 University Road, HsinChu 30010, Taiwan
paulachen@g2.nctu.edu.tw

Haiyan Wu
Department of Educational Psychology and Learning Systems
Florida State University
Tallahassee, FL 32304, USA
hw07d@fsu.edu