The Family of Log-Skew-Normal Alpha-Power Distributions using Precipitation Data

La familia de distribuciones alfa-potencia log-skew-normal usando datos de precipitación

Guillermo Martínez-Flórez[1,a], Sandra Vergara-Cardozo[2,b], Luz Mery González[2,c]

[1]Departamento de Matemáticas y Estadística, Universidad de Córdoba, Montería, Colombia

[2]Departamento de Estadística, Facultad de Ciencias, Univesidad Nacional de Colombia, Bogotá D.C, Colombia

---

Abstract

We present a new set of distributions for positive data based on a skew-normal alpha-power (PSN) model including a new parameter which in turn makes the log-skew-normal alpha-power (LPSN) model more flexible than both the log-normal (LN) model and log-skew-normal (LSN) model. The LPSN model contains the LN model and LSN model as special cases. Furthermore, it models positive data with asymmetry and kurtosis larger than the one permitted by the LN distribution. Precipitation data illustrates the usefulness of the LPSN model being less influenced by outliers.

Key words: Asymmetry, Fisher information matrix, Kurtosis, Likelihood ratio test, Maximum likelihood estimator.

Resumen

Presentamos una nueva familia de distribuciones para datos positivos basada en el modelo skew-normal alpha-power (PSN), incluyendo un nuevo parámetro el cual hace el modelo log-skew-normal alpha-power (LPSN) más flexible que los modelos log-normal (LN) y log-skew-normal (LSN). El modelo LPSN contiene el modelo LN y el modelo LSN como casos particulares. Además, modela datos positivos con asimetría y curtosis más allá de lo permitido por la distribución LN. Datos de precipitación ilustran la utilidad del modelo LPSN siendo menos influenciado por outliers.

Palabras clave: asimetría, curtosis, estimador máxima verosimilitud, matriz de información de Fisher, test de razón de verosimilitud.

[a]Professor. E-mail: gmartinez@correo.unicordoba.edu.co

[b]Assistant professor. E-mail: svergarac@unal.edu.co

[c]Assistant professor. E-mail: lgonzalezg@unal.edu.co

1. Introduction

The log-normal (LN) distribution obtained as a transformation of the normal distribution has been widely used to model different types of information including income in economics and material lifetimes. In of different fields of knowledge, asymmetry and kurtosis of the data are outside of the range allowed by the LN distribution so it is necessary to use another distribution that can take into account these issues. In the same way that Azzalini (1985), we introduce the skew-normal (SN) distribution to conform data with a range of asymmetry and kurtosis outside the range allowed by the normal distribution, Lin & Stoyanov (2009) present the log-skew-normal (LSN) distribution which is an extension for positive data of the LN distribution in order to conform data with asymmetry and kurtosis outside the range allowed by the LN distribution. The probability density function of this model is given by

$$
\begin{aligned}
\varphi_{LSN}(y; \xi, \eta, \lambda) &= \frac{2}{\eta y} \phi\left(\frac{\log(y) - \xi}{\eta}\right) \left\{\Phi\left(\lambda \frac{\log(y) - \xi}{\eta}\right)\right\} \\
&= \frac{1}{y} \phi_{SN}\left(\log(y); \xi, \eta, \lambda\right), \ y \in \mathbb{R}^+
\end{aligned}
\tag{1}
$$

where

$$
\phi_{SN}(x; \xi, \eta, \lambda) = \frac{2}{\eta} \phi\left(\frac{x - \xi}{\eta}\right) \left\{\Phi\left(\lambda \frac{x - \xi}{\eta}\right)\right\}
$$

denotes the density function of the SN distribution with parameters of location (ξ), scale (η), and shape (λ). The LSN model $[Y \sim LSN(\xi, \eta, \lambda)]$ given by (1) contains the parameters of location (ξ), scale (η), and shape (λ) that control the asymmetry of the data. $\phi(.)$ and $\Phi(.)$ denote the density and cumulative distribution function of standard normal distribution, N(0,1). Based on the SN of Azzalini (1985) and generalized Gaussian (PN) of Durrans (1992), Martínez-Flórez (2011) introduce and studies the main features of the asymmetric distribution called skew-normal alpha-power (PSN) distribution with probability density function given by

$$
\phi_{PSN}(z; \lambda, \alpha) = \alpha \phi_{SN}(z; \lambda) \left\{\Phi_{SN}(z; \lambda)\right\}^{\alpha - 1}
\tag{2}
$$

where $z, \lambda \in \mathbb{R}$, $\alpha \in \mathbb{R}^+$, $\phi_{SN}(z; \lambda) = \phi_{SN}(z; 0, 1, \lambda)$ as defined in (1) and $\Phi_{SN}(z; \lambda)$ in (3). The PSN model $[X \sim PSN(\lambda, \alpha)]$ given by (2) considers parameters of shape λ and α with

$$
\Phi_{SN}(z; \lambda) = \int_{-\infty}^{z} \phi_{SN}(t; \lambda) dt = \Phi(z) - 2T(z, \lambda)
\tag{3}
$$

being the cumulative distribution function of skew-normal distribution, Azzalini (1985), and $T(., \lambda)$ the Owen's (1956) function.

In (2), $\lambda = 0$ and $\alpha = 1$ corresponds to the standard normal case, *i.e.*, $\phi_{SN}(.; 0, 1, 0) = \phi_{PSN}(z; 0, 1) = \phi_{SN}(.; 0) = \phi(.)$ and $\Phi_{SN}(.; 0) = \Phi(.)$. The model is an extension of the PN model, Durrans (1992) and the Gupta & Gupta (2008) exponential model

$$\varphi_\alpha(z;\alpha) = \alpha\phi(z)\{\Phi(z)\}^{\alpha-1}, \quad z \in \mathbb{R} \tag{4}$$

replacing the normal density by the skew-normal density.

Martínez-Flórez (2011) demonstrate that the expected information matrix of the PSN model is nonsingular in the neighborhood of the skewness parameters $\lambda = 0$ and $\alpha = 1$ contrary to the case of Azzalini (1985) whose expected information matrix is singular in the neighborhood of $\lambda = 0$. Table 1 shows the intervals of asymmetry and kurtosis coefficients for the PSN, SN, and PN models. The PSN model has greater asymmetry and the distribution is more platikurtic or leptokurtic than the Azzalini (1985) and Durrans (1992) models. This shown an advantage of the model (2) over the $\phi_{SN}(z;\lambda)$ and $\varphi_\alpha(z;\alpha)$ models.

TABLE 1: Intervals of asymmetry ($\sqrt{\beta_1}$) and kurtosis (β_2) coefficients, defined in (7), for the PSN, SN, and PN models given by Martínez-Flórez, G. (2011).

Model	$\sqrt{\beta_1}$	β_2
Skew-normal alpha-power (PSN) model	[-1.4676 ; 0.9953)	[1.4672 ; 5.4386]
Skew-normal (SN) model	(-0.9953 ; 0.9953)	[3 ; 3.8692)
Generalized gaussian (PN) model	[-0.6115 ; 0.9007]	[1.7170 ; 4.3556]

Figures 1(a) show corresponding and 1(b), the parameters λ and α of asymmetry and kurtosis of the PSN distribution a more flexible model than Azzalini (1985) and Durrans (1992) yielding.
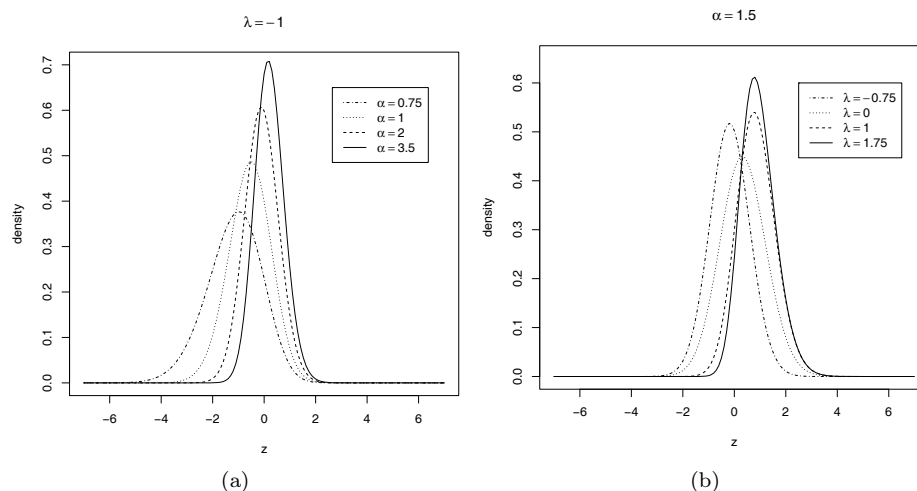

FIGURE 1: Probability density function of the skew-normal alpha-power distribution.

Other work on this type of distribution was studied by Arnold & Beaver (2002) and Gupta & Gupta (2004). We present a new set of distributions based on the PSN distribution that corresponds to the log-skew-normal alpha-power (LPSN) distribution.

In Section 2, we describe the LPSN distribution, its observed information matrix and the expected information matrix. We also perform an application of the proposed model to data by IDEAM (2006) in which the coefficients of skewness and kurtosis of the model justify the use of the LPSN model. We conclude with a brief discussion in Section 3.

2. Log-Skew-Normal Alpha-Power (LPSN) Model

The LPSN distribution is a new alternative for family distribution of positive data with a range of asymmetry and/or kurtosis outside of the range permitted by the LN and LSN distributions.

Definition 1. The positive random variable Y in the $\mathbb{R}^+$ has a univariate log-skew-normal alpha-power distribution with parameters λ and α if the transformed variable $Z = \log(Y)$ has a PSN distribution with parameters λ and α. This is denoted by $Y \sim LPSN(\lambda, \alpha)$. The probability density function of a random variable Y with distribution $LPSN(\lambda, \alpha)$ is given by

$$\varphi_{LPSN}(y; \lambda, \alpha) = \frac{\alpha}{y} \phi_{SN}(\log(y); \lambda) \left\{ \Phi_{SN}(log(y); \lambda) \right\}^{\alpha-1}, \ y, \ \alpha \in \mathbb{R}^+ \text{ and } \lambda \in \mathbb{R}$$

The cumulative distribution function of the LPSN model is given by

$$\mathcal{F}_Y(y; \lambda, \alpha) = \left\{ \Phi_{SN}(log(y); \lambda) \right\}^{\alpha}, \ y \in \mathbb{R}^+ \tag{5}$$

According to equation (5), the inversion method can be used to generate a random variable with distribution $LPSN(\lambda, \alpha)$. Thus, if U is a uniform random variable in (0,1) the random variable $Y = exp\{\Phi_{ISN}(U^{1/\alpha}; \lambda)\}$ has LPSN distribution of the parameters λ and α where Φ_{ISN} represents the inverse function of the SN distribution, $\Phi_{SN}(.; \lambda)$, whose values can be obtained in many statistical packages (R Development Core Team 2011).

When $\alpha = 1$, the LPSN distribution is identical to the LSN distribution $[\varphi_{LPSN}(y; \lambda, 1) = \varphi_{LSN}(y; 0, 1, \lambda)]$ and when $\lambda = 0$ and $\alpha = 1$, the LPSN distribution is identical to the log-normal (LN) distribution. So LPSN distribution is more flexible than LN and LSN distributions (see, for example, Figures 2(a) and 2(b)).

2.1. Moments of the Distribution

The *r-th* moment of the random variable Y with LPSN distribution can be written as,

$$\mu_r = \mathbb{E}(Y^r) = \alpha \int_0^1 \left\{ \exp\left[r\Phi_{ISN}(y; \lambda) \right] \right\} y^{\alpha-1} dy \tag{6}$$

Let $\mu_r' = \mathbb{E}(Y - \mathbb{E}(Y))^r$, $r = 2, 3, 4$,

$$\mu_2' = \mu_2 - \mu_1^2, \ \ \mu_3' = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \ \text{ and } \ \mu_4' = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4$$
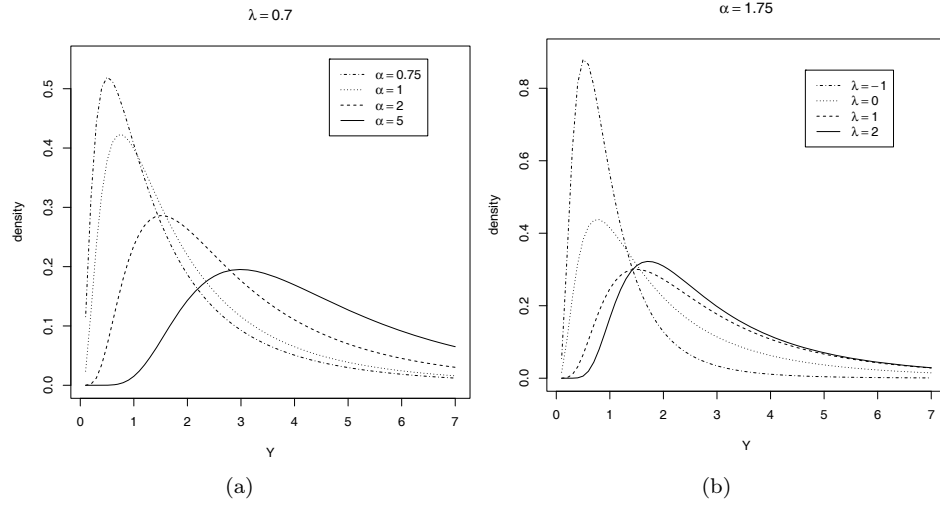
FIGURE 2: Probability density function of the log-skew-normal alpha-power distribution.

The variance, coefficient of variation, skewness and kurtosis are given by:

$$\sigma_Y^2 = Var(Y) = \mu_2', \quad CV = \frac{\sqrt{\sigma_Y^2}}{\mu_1}, \quad \sqrt{\beta_1} = \frac{\mu_3'}{[\mu_2']^{3/2}} \quad \text{and} \quad \beta_2 = \frac{\mu_4'}{[\mu_2']^2} \qquad (7)$$

2.2. Scale-Location

Let $PSN(\xi, \eta, \lambda, \alpha)$ denotes a location-scale transformation of $PSN(\lambda, \alpha)$ where $\xi \in \mathbb{R}$, $\eta \in \mathbb{R}^+$ and $Y = \xi + \eta Z$.

Definition 2. If X has a distribution of localization-scale parameters $PSN(\xi, \eta, \lambda, \alpha)$ then the extension of scale-location to the LPSN distribution follows the transformation $X = \log(Y)$, where $\xi \in \mathbb{R}$ and $\eta \in \mathbb{R}^+$. Then, the density of Y is given by

$$\varphi_{LPSN}(y; \xi, \eta, \lambda, \alpha) = \alpha \varphi_{LSN}(y; \xi, \eta, \lambda) \left\{ \Phi_{SN} \left(\frac{\log(y) - \xi}{\eta}; \lambda \right) \right\}^{\alpha-1} \qquad (8)$$

$$y, \ \alpha \in \mathbb{R}^+, \text{ and } \lambda \in \mathbb{R}$$

where $\varphi_{LSN}(y; \xi, \eta, \lambda)$ is defined in (1) and $\Phi_{SN}(.; \lambda)$, in (3)

We use the notation $Y \sim LPSN(\xi, \eta, \lambda, \alpha)$. So $LPSN(\lambda, \alpha) = LPSN(0, 1, \lambda, \alpha)$.

A special case in the model (8) is when $\lambda = 0$, obtaining the density,

$$\varphi_{LPSN}(y; \xi, \eta, 0, \alpha) = \frac{\alpha}{\eta y} \phi \left(\frac{\log(y) - \xi}{\eta} \right) \left\{ \Phi \left(\frac{\log(y) - \xi}{\eta} \right) \right\}^{\alpha-1}, \ y \in \mathbb{R}^+$$

This is denoted $Y \sim LPSN_{\lambda=0}(\xi, \eta, \alpha)$. Like the model LSN model, this distribution is also a generalization of the LN model which we will call the generalized LN distribution.

The following result is an extension of the LN and LSN distributions.

Theorem 1. *For any $\lambda \in \mathbb{R}$ and $\alpha \in \mathbb{R}^+$, the random variable $Y \sim LPSN(\xi, \eta, \lambda, \alpha)$ does not have a moment generating function (MGF).*

Proof. As $\lambda = 0$ and $\alpha = 1$ in the LPSN model, we have the case of the LN distribution, which does not have a moment generating function. Since MGF satisfies the property,
$$M_{aY+b}(t) = \exp(bt)M_Y(at)$$
then it is sufficient to consider the standard case $LPSN(\lambda, \alpha)$. $\qquad \square$

For fixed values $\alpha = \alpha_0 > 0$ and $\lambda = \lambda_0$, the MGF of Y can be written as

$$\begin{aligned}
M_Y(t) &= \mathbb{E}(e^{ty}) \\
&= \int_0^\infty e^{ty} \varphi_{LPSN}(y; \lambda_0, \alpha_0) dy \\
&= \int_0^\infty \frac{\alpha_0}{y} e^{ty} \phi_{SN}(log(y); \lambda_0) \{\Phi_{SN}(log(y); \lambda_0)\}^{\alpha_0 - 1} dy \\
&= \int_0^\infty h(y, t, \lambda_0, \alpha_0) g(y, \lambda_0, \alpha_0) dy, \quad y \in \mathbb{R}^+
\end{aligned}$$

with

$$h(y, t, \lambda_0, \alpha_0) = \frac{2\alpha_0}{y} e^{ty} \phi(\log(y))\{\Phi(\lambda_0 \log(y))\} > 0$$

and

$$g(y, \lambda_0, \alpha_0) = \{\Phi_{SN}(log(y); \lambda_0)\}^{\alpha_0 - 1}$$

to all $y > 0$.

When $t > 0$ is fixed, we prove that

$$J_{(\lambda_0, \alpha_0)} = \int_0^\infty h(y, t, \lambda_0, \alpha_0) g(y, \lambda_0, \alpha_0) dy = \infty$$

for all $\lambda_0 \in \mathbb{R}$ and $\alpha_0 \in \mathbb{R}^+$.

If $\lambda_0 > 0$ according to Lin & Stoyanov (2009)

$$\liminf_{y \to \infty} \{\Phi(\lambda_0 \log(y))\} \geq \frac{1}{2}$$

therefore $h(y, t, \lambda_0, \alpha_0) \to \infty$ when $y \to \infty$. Now, $g(y, \lambda_0, \alpha_0) \to 1$ when $y \to \infty$, then we conclude that $J_{(\lambda_0, \alpha_0)} \to \infty$ when $y \to \infty$.

According to Lin & Stoyanov (2009), if $\lambda_0 < 0$ then

$$\lim_{y \to \infty} \frac{-\log(\Phi(-y))}{y^2} = \frac{1}{2}$$

Therefore, when $y \to \infty$, we have the asymptotic approximation,

$$\log\left(\Phi(\lambda_0 \log(y))\right) \approx \frac{1}{2}\left(\lambda_0 \log(y)\right)^2$$

Then, we assume that $log(\alpha_0) < \infty$, where $y \to \infty$ must be

$$\log\left(h(y,t,\lambda_0,\alpha_0)\right) - \log(\alpha_0) \approx \frac{1}{2}\log\left(\frac{2}{\pi}\right) - \log(y) + ty - \frac{1}{2}(\lambda_0^2 + 1)(\log(y))^2 \to \infty$$

Now, since $g(y,\lambda_0,\alpha_0) \to 1$, when $y \to \infty$, then we conclude that $J_{(\lambda_0,\alpha_0)} \to \infty$ when $y \to \infty$.

2.3. Inference

The maximum likelihood estimation and observed and expected matrix information for the parameters of the $LPSN(\xi,\eta,\lambda,\alpha)$ model are studied. For a random sample of size n, $Y_1, Y_2, \ldots, Y_n$, with $Y_i \sim LPSN(\xi,\eta,\lambda,\alpha)$, the log-likelihood function of $\theta = (\xi,\eta,\lambda,\alpha)'$ given $\mathbf{Y}$, can be expressed by

$$\ell(\theta,\mathbf{Y}) = n\left(\log(\alpha) - \log(\eta)\right) - \sum_{i=1}^{n}\log(y_i) - \frac{1}{2}\sum_{i=1}^{n}z_i^2$$
$$+ \sum_{i=1}^{n}\log\left\{\Phi(\lambda z_i)\right\} + (\alpha - 1)\sum_{i=1}^{n}\log\left\{\Phi_{SN}(z_i;\lambda)\right\}$$

where $z_i = \frac{\log(y_i) - \xi}{\eta}$. The elements of the score function are given by

$$U(\xi) = \frac{1}{\eta}\sum_{i=1}^{n}z_i - \frac{\lambda}{\eta}\sum_{i=1}^{n}w_i - \frac{\alpha - 1}{\eta}\sum_{i=1}^{n}w_{1i}$$

$$U(\eta) = -\frac{n}{\eta} + \frac{1}{\eta}\sum_{i=1}^{n}z_i^2 - \frac{\lambda}{\eta}\sum_{i=1}^{n}z_i w_i - \frac{\alpha - 1}{\eta}\sum_{i=1}^{n}w_{1i}z_i$$

$$U(\lambda) = \sum_{i=1}^{n}z_i w_i - \sqrt{\frac{2}{\pi}}\frac{(\alpha - 1)}{1 + \lambda^2}\sum_{i=1}^{n}w_i(\lambda)$$

and

$$U(\alpha) = \frac{n}{\alpha} + \sum_{i=1}^{n}\log\left\{\Phi_{SN}(z_i;\lambda)\right\}$$

where $w = \frac{\phi(\lambda z)}{\Phi(\lambda z)}$, $w_1 = \frac{\phi_{SN}(z)}{\Phi_{SN}(z;\lambda)}$ and $w(\lambda) = \frac{\phi\left(\sqrt{1+\lambda^2}z\right)}{\Phi_{SN}(z;\lambda)}$. The score equations are obtained by equating these partial derivatives to zero. The maximum likelihood estimators (MLEs) are the solutions to the score equations. These solutions are usually obtained by iterative numerical methods.

2.3.1. Observed Information Matrix

The elements of the observed information matrix are defined without the second derivative of the log-likelihood function with respect to parameter denoted by $j_{\xi\xi}, j_{\eta\xi}, \ldots, j_{\alpha\alpha}$ which can be written as

$$j_{\xi\xi} = \frac{n}{\eta^2} + \frac{\lambda^2}{\eta^2} \sum_{i=1}^{n} \lambda z_i w_i + \frac{\lambda^2}{\eta^2} \sum_{i=1}^{n} w_i^2$$
$$+ \frac{\alpha-1}{\eta^2} \sum_{i=1}^{n} w_{1i}(z_i + w_{1i}) - \sqrt{\frac{2}{\pi}} \frac{\lambda(\alpha-1)}{\eta^2} \sum_{i=1}^{n} w_i(\lambda)$$

$$j_{\eta\xi} = \frac{2}{\eta^2} \sum_{i=1}^{n} z_i + \frac{\lambda^3}{\eta^2} \sum_{i=1}^{n} z_i^2 w_i + \frac{\lambda^2}{\eta^2} \sum_{i=1}^{n} z_i w_i^2 - \frac{\lambda}{\eta^2} \sum_{i=1}^{n} w_i$$
$$- \sqrt{\frac{2}{\pi}} \frac{\lambda(\alpha-1)}{\eta^2} \sum_{i=1}^{n} z_i w_i(\lambda) + \frac{\alpha-1}{\eta^2} \sum_{i=1}^{n} w_{1i}(-1 + z_i^2 + z_i w_{1i})$$

$$j_{\lambda\xi} = \frac{1}{\eta} \sum_{i=1}^{n} \left[w_i - \lambda^2 z_i^2 w_i - \lambda z_i w_i^2 \right]$$
$$+ \sqrt{\frac{2}{\pi}} \frac{\alpha-1}{\eta} \sum_{i=1}^{n} w_i(\lambda) \left[z_i + \frac{1}{1+\lambda^2} w_{1i} \right], \quad j_{\alpha\xi} = \frac{1}{\eta} \sum_{i=1}^{n} w_{1i}$$

$$j_{\eta\eta} = -\frac{n}{\eta^2} + \frac{3}{\eta^2} \sum_{i=1}^{n} z_i^2 - \frac{2\lambda}{\eta^2} \sum_{i=1}^{n} z_i w_i + \frac{\lambda^3}{\eta^2} \sum_{i=1}^{n} z_i^3 w_i + \frac{\lambda^2}{\eta^2} \sum_{i=1}^{n} z_i^2 w_i^2$$
$$- \sqrt{\frac{2}{\pi}} \frac{\lambda(\alpha-1)}{\eta^2} \sum_{i=1}^{n} z_i^2 w_i(\lambda) + \frac{\alpha-1}{\eta^2} \sum_{i=1}^{n} z_i w_{1i} \left[-2 + z_i^2 + z_i w_{1i} \right]$$

$$j_{\lambda\eta} = \frac{1}{\eta} \sum_{i=1}^{n} z_i w_i - \frac{\lambda^2}{\eta} \sum_{i=1}^{n} z_i^3 w_i - \frac{\lambda}{\eta} \sum_{i=1}^{n} z_i^2 w_i^2$$
$$+ \sqrt{\frac{2}{\pi}} \frac{\alpha-1}{\eta} \sum_{i=1}^{n} z_i w_i(\lambda) \left[z_i + \frac{1}{1+\lambda^2} w_{1i} \right]$$

$$j_{\lambda\lambda} = \sum_{i=1}^{n} z_i^2 (\lambda z_i w_i + w_i^2) - \sqrt{\frac{2}{\pi}} \frac{2\lambda(\alpha-1)}{(1+\lambda^2)^2} \sum_{i=1}^{n} w_i(\lambda)$$
$$+ 2(\alpha-1) \sum_{i=1}^{n} \left[-\sqrt{\frac{1}{2\pi}} \frac{\lambda}{1+\lambda^2} z_i^2 w_i(\lambda) + \frac{1}{\pi} \frac{1}{(1+\lambda^2)^2} w_i^2(\lambda) \right]$$

and

$$j_{\alpha\eta} = \frac{1}{\eta} \sum_{i=1}^{n} z_i w_{1i}, \qquad j_{\alpha\lambda} = \sqrt{\frac{2}{\pi}} \frac{1}{1+\lambda^2} \sum_{i=1}^{n} w_i(\lambda), \qquad j_{\alpha\alpha} = \frac{n}{\alpha^2}$$

2.3.2. Expected Information Matrix

The elements of the expected information matrix are the expected values of the elements of the observed information matrix; let $i_{\xi\xi}, i_{\eta\xi}, \ldots, i_{\alpha\alpha}$ be the elements of the observed information matrix multiplied by n^{-1}, calling $a_{jk} = \mathbb{E}(z^j w^k)$, $a_{1jk} = \mathbb{E}(z^j w_1^k)$ and $a_{jk}(\lambda) = \mathbb{E}(z^j w^k(\lambda))$. The elements of the expected information matrix can be written as

$$i_{\xi\xi} = \frac{1}{\eta^2} + \frac{\lambda^3}{\eta^2} a_{11} + \frac{\lambda^2}{\eta^2} a_{02} - \sqrt{\frac{2}{\pi}} \frac{\lambda(\alpha-1)}{\eta^2} a_{01}(\lambda) + \frac{\alpha-1}{\eta^2}(a_{111} + a_{102})$$

$$i_{\eta\xi} = \frac{2}{\eta^2} a_{10} + \frac{\lambda^3}{\eta^2} a_{21} + \frac{\lambda^2}{\eta^2} a_{12} - \frac{\lambda}{\eta^2} a_{10} - \sqrt{\frac{2}{\pi}} \frac{\lambda(\alpha-1)}{\eta^2} a_{11}(\lambda)$$
$$+ \frac{\alpha-1}{\eta^2}(-a_{101} + a_{121} + a_{112})$$

$$i_{\lambda\xi} = \frac{1}{\eta} \left[a_{01} - \lambda^2 a_{21} - \lambda a_{12} \right] + \sqrt{\frac{2}{\pi}} \frac{\alpha-1}{\eta} \left[a_{11}(\lambda) \right.$$
$$\left. + \frac{1}{1+\lambda^2} \mathbb{E}(w_1 w(\lambda)) \right], \quad i_{\alpha\xi} = \frac{1}{\eta} a_{101}$$

$$i_{\eta\eta} = -\frac{1}{\eta^2} + \frac{3}{\eta} a_{20} - \frac{2\lambda}{\eta^2} a_{11} + \frac{\lambda^3}{\eta} a_{31} + \frac{\lambda^2}{\eta^2} a_{22} - \sqrt{\frac{2}{\pi}} \frac{\lambda(\alpha-1)}{\eta^2} a_{21}(\lambda)$$
$$+ \frac{\alpha-1}{\eta^2}(-2a_{111} + a_{131} + a_{122})$$

$$i_{\lambda\eta} = \frac{1}{\eta} \left[a_{11} - \lambda^2 a_{31} - \lambda a_{22} \right] + \sqrt{\frac{2}{\pi}} \frac{\alpha-1}{\eta} \left[a_{21}(\lambda) + \right.$$
$$\left. \frac{1}{1+\lambda^2} \mathbb{E}(z w_1 w(\lambda)) \right], \quad i_{\alpha\eta} = \frac{1}{\eta} a_{111}$$

$$i_{\lambda\lambda} = \lambda a_{31} + a_{22} - \sqrt{\frac{2}{\pi}}\frac{2\lambda(\alpha - 1)}{(1+\lambda^2)^2}a_{01}(\lambda) + \sqrt{\frac{2}{\pi}}(\alpha - 1)\Big[-\frac{\lambda}{1+\lambda^2}a_{21}(\lambda)$$

$$+ \sqrt{\frac{2}{\pi}}\frac{1}{(1+\lambda^2)^2}a_{02}(\lambda)\Big]$$

$$i_{\alpha\lambda} = \sqrt{\frac{2}{\pi}}\frac{1}{1+\lambda^2}a_{01}(\lambda), \qquad i_{\alpha\alpha} = \frac{1}{\alpha^2}$$

For $\lambda = 0$ and $\alpha = 1$ use the approximation

$$\frac{1}{\pi}\frac{\phi(z)}{\sqrt{\Phi(z)[1-\Phi(z)]}} \approx \frac{1}{\sqrt{2\pi}(\pi/2)}\exp\left(-\frac{z^2}{2(\pi^2/4)}\right)$$

given in Chaibub-Neto & Branco (2003). The expected information matrix is

$$I_F(\theta) = \begin{pmatrix} \frac{1}{\eta^2} & 0 & \sqrt{\frac{2}{\pi}}\frac{1}{\eta} & \frac{\sqrt{\pi}}{2}\frac{1}{\eta} \\ 0 & \frac{2}{\eta^2} & 0 & \frac{1}{4\eta}\frac{\pi^2}{\sqrt{8+\pi^2}} \\ \sqrt{\frac{2}{\pi}}\frac{1}{\eta} & 0 & \frac{2}{\pi} & \sqrt{\frac{1}{2}} \\ \frac{\sqrt{\pi}}{2}\frac{1}{\eta} & \frac{1}{4\eta}\frac{\pi^2}{\sqrt{8+\pi^2}} & \sqrt{\frac{1}{2}} & 1 \end{pmatrix} \qquad (9)$$

whose determinant $|I_F(\theta)| = 0$.

Therefore, we conclude that the expected information matrix of the model is singular for the special case of a LN distribution. The upper 3×3 submatrix is the expected information matrix from the log-skew-normal distribution.

As in (9) the third column (respectively, row) is equal to first column (respectively, row) multiply by $\eta\sqrt{\frac{2}{\pi}}$, $I_F(\theta)$ is singular. Using results from Rotnitzky, Cox, Bottai & Robins (2000) we find the asymptotic distribution of the maximum likelihood estimator of θ. DiCiccio & Monti (2004) explains: "(Rotnitzky et al. 2000) *derived the asymptotic distribution of the MLE* $\widehat{\theta} = (\widehat{\theta}^1, \widehat{\theta}^2, \ldots, \widehat{\theta}^q)$ *under two conditions: a single component of the score function, say* S_{θ^1}, *vanishes at some point* $\theta = \theta^*$, *and some higher-order derivatives of* S_{θ^1} *taken with respect to* θ^1 *are possibly 0 at that point but the first nonzero derivative is not a linear combination of the other score function components* $S_{\theta^2}, \ldots, S_{\theta^q}$".

Using an iterative process suggested by Rotnitzky et al. (2000), we find a new parameterization to PSN model that fulfill the two conditions in the same way that Chiogna (1998) and DiCiccio & Monti (2004) for the skew-normal distribution and the skew exponential power distribution, respectively. Let $\theta* = (\xi*, \eta*, 0, 1)$ denote the vector parameter of interest. For $\theta = \theta*$, let $S_\theta(\theta*, Y) = \partial\ell/\partial\theta* = (S_\xi^*, S_\eta^*, S_\lambda^*, S_\alpha^*)$ denote the score vector, so

$$S_\theta(\theta*, Y) = \left(\frac{Z^*}{\eta^*}, \frac{Z^{*2} - 1}{\eta^*}, \sqrt{\frac{2}{\pi}}Z^*, 1 + \log(\Phi(Z^*))\right)$$

whit $Z^* = \frac{Y - \xi^*}{\eta^*}$. After some calculations we take the new parameterization $\tilde{\theta} = \tilde{\theta}(\theta) = (\tilde{\xi}, \tilde{\eta}, \lambda, \alpha)$ with $\tilde{\xi} = \xi + \sqrt{\frac{2}{\pi}}\eta^*\lambda$ and $\tilde{\eta} = \eta - \eta^*\frac{\lambda^2}{\pi}$.

Making use of Theorem 3 in Rotnitzky et al. (2000) with the new parameterization we can conclude that:

1. The MLE of θ is unique with probability tending to 1, and it is consistent.

2. The likelihood ratio statistic for testing the simple null hypothesis $H_0 : \theta = \theta^*$ converges in distribution to the χ^2 distribution with four degrees of freedom.

3. The random vector
$$\left(n^{1/2}(\tilde{\xi} - \xi + \sqrt{\tfrac{2}{\pi}}\eta^*\lambda), n^{1/2}(\tilde{\eta} - \eta - \eta^*\tfrac{\lambda^2}{\pi}), n^{1/6}\widehat{\lambda}, n^{1/2}(\widehat{\alpha} - 1) \right)$$

converges to $(Y_1, Y_2, Y_3^{1/3}, Y_4)$, where (Y_1, Y_2, Y_3, Y_4) is a normal random vector with mean zero and covariance matrix equals to the inverse of the covariance matrix

$$\begin{pmatrix} \frac{1}{\eta^2} & 0 & \frac{2-\pi}{\sqrt{2\pi^3}}\frac{1}{\eta} & \frac{\sqrt{\pi}}{2}\frac{1}{\eta} \\ 0 & \frac{2}{\eta^2} & 0 & \frac{1}{4\eta}\frac{\pi^2}{\sqrt{8+\pi^2}} \\ \frac{2-\pi}{\sqrt{2\pi^3}}\frac{1}{\eta} & 0 & \frac{5\pi^2-28\pi+44}{6\pi^3} & \sqrt{\frac{1}{2}} \\ \frac{\sqrt{\pi}}{2}\frac{1}{\eta} & \frac{1}{4\eta}\frac{\pi^2}{\sqrt{8+\pi^2}} & \sqrt{\frac{1}{2}} & 1 \end{pmatrix}^{-1}$$

2.4. Illustration

Precipitation data (measured in inches) were collected from the Colombian Institute of Hydrology, Meteorology and Environmental Studies in Córdoba, Colombia (IDEAM 2006). Descriptive statistics for the variable under study are provided in Table 2. The quantities $\sqrt{\widehat{\beta_1}} = \sqrt{b_1}$ and $\widehat{\beta_2} = b_2$, where β_1 and β_2 defined in (7), indicate the asymmetry and kurtosis coefficients respectively.

TABLE 2: Descriptive statistics of the precipitation variable

Variables	n	Mean	Variance	$\sqrt{b_1}$	b_2
Y	273	4.8360	9.7871	0.4632	2.6035
$log(Y)$	273	1.2219	1.1155	-1.5608	5.5276

The asymmetry and kurtosis coefficients are different from the corresponding values expected for LN model and normal model. Precipitation data are fitted using the LPSN model.

The LPSN model is compared to the LN model as well as the LSN model to the $LPSN_{\lambda=0}$ model. The maximum likelihood method for estimating the parameters is used and the Akaike information criterion (AIC), (Akaike 1974), is applied for

contrast. Firstly, the LN model is compared to the LPSN model by the hypothesis tests

$$H_0 : (\lambda, \alpha) = (0, 1) \quad \text{versus} \quad H_1 : (\lambda, \alpha) \neq (0, 1)$$

Using the likelihood ratio statistic,

$$\Lambda = \frac{\ell_{LN}(\widehat{\theta})}{\ell_{LPSN}(\widehat{\theta})}$$

we obtain

$$-2\log(\Lambda) = -2(-735.4023 + 670.2293) = 130.346$$

which is greater than the value of the $\chi^2_{2,95\%} = 5.99$. Then the LPSN model is a good alternative for fitting the precipitation data. The LPSN model is also compared to the $LPSN_{\lambda=0}$ model and the LSN models by the hypothesis tests

$$H_{01} : \lambda = 0 \quad \text{versus} \quad H_{11} : \lambda \neq 0, \quad \text{and} \quad H_{02} : \alpha = 1 \quad \text{versus} \quad H_{12} : \alpha \neq 1$$

respectively, using the likelihood ratio statistics

$$\Lambda_1 = \frac{\ell_{LPSN_{\lambda=0}}(\theta)}{\ell_{LPSN}(\theta)} \quad \text{and} \quad \Lambda_2 = \frac{\ell_{LSN}(\theta)}{\ell_{LPSN}(\theta)}$$

After numerical evaluations, we obtain

$$-2\log(\Lambda_1) = 61.5960 \quad \text{and} \quad -2\log(\Lambda_2) = 15.5056$$

which is greater than the value of the $\chi^2_{1,95\%} = 3.84$. The best fit, with respect to the other models, is shown by the LPSN model. Table 3 presents the MLEs and the estimated standard errors (in parentheses) for LN, LSN, LPSN and models. Figure 3 shows the histogram of precipitacion data and fitted curves for the proposed models in which the LPSN model presents the better fit of asymmetry and kurtosis with respect to the other models.

TABLE 3: Parameters and estimated standard errors of the log-normal (LN), log-skew-normal (LSN), log-skew-normal alpha-power $\lambda = 0$ ($LPSN_{\lambda=0}$), and the log-skew-normal alpha-power (LPSN) distributions.

Parameter	Log-normal	LSN	LPSN$_{\lambda=0}$	LPSN
$Loglik$	-735.4023	-677.9821	-701.0273	-670.2292
AIC	1474.8050	1361.964	1408.0550	1348.5490
ξ	1.2219(0.0638)	2.4217(0.0392)	2.8280(0.0817)	2.2647(0.0529)
η	1.0542(0.0451)	1.5971(0.0763)	0.1668(0.0507)	4.8760(0.3363)
λ	–	-10.0515(2.2917)	–	-19.2702(2.4450)
α	–	–	0.0144(0.0008)	4.8579(0.5925)

The Figure 4 shows the qqplots for LN, LSN and LPSN models. The LPSN model shows better fit with respect to the LN and LSN models.
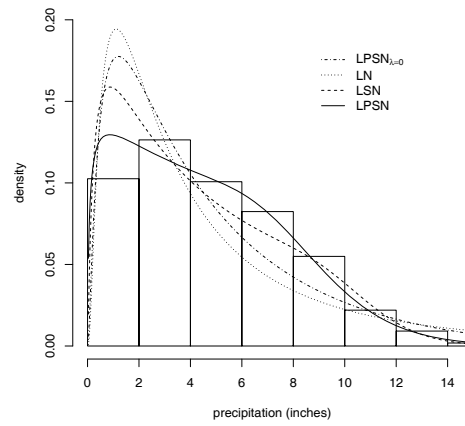
FIGURE 3: Histogram of the precipitation data. Densities are estimated by maximum likelihood.
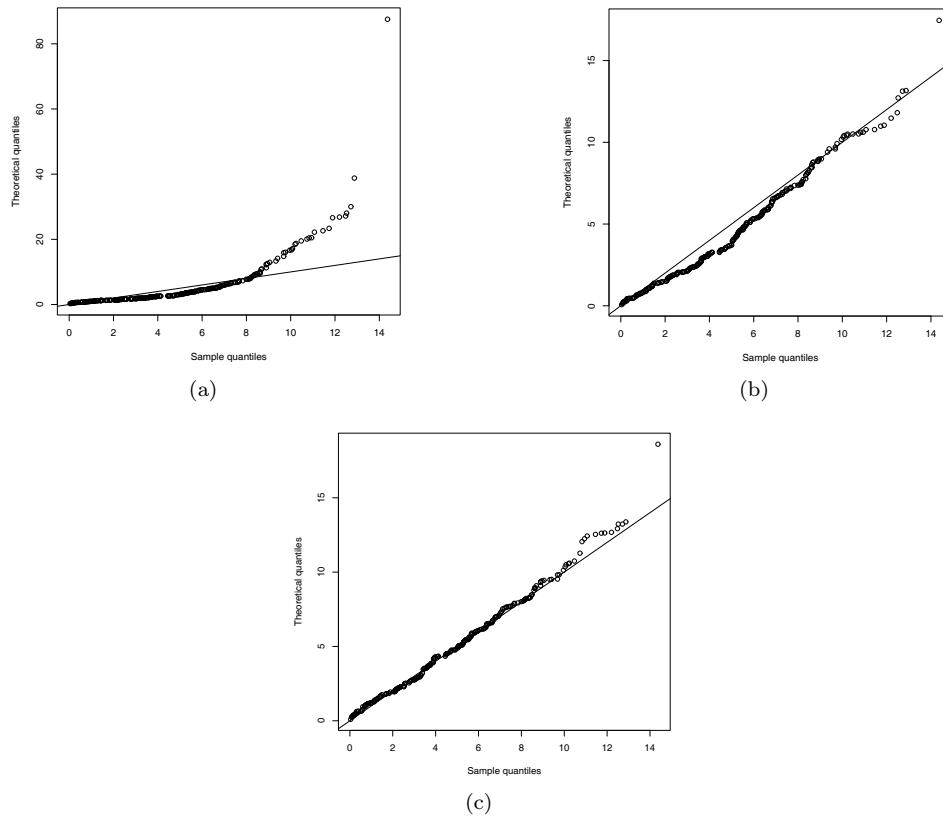

(a)

(b)

(c)

FIGURE 4: Q-Qplot: (a) log-normal model, (b) log-skew-normal model, and (c) log-skew-normal alpha-power model.

3. Conclusion

In this paper we propose a more flexible model than LN and LSN models
fit data with greater asymmetry and more platikurtic or leptokurtic than Azzalini
(1985) and Durrans (1992) models. General expressions for the moments are found,
maximum likelihood estimators are studied, observed and expected information
matrix are found, and also an asymptotic distribution of a MLEs vector is found.
Finally, an illustration is presented (see Figure 4). We contrast the LN, LSN,
and LPSN models through some precipitation data. According to AIC selection
criterion, the LPSN model makes the better fit with respect to the other models
considered.

References

Akaike, H. (1974), 'A new look at statistical model identification', *IEEE Transaction on Automatic Control* **Ac-19**(9), 716–722.

Arnold, B. C. & Beaver, R. (2002), 'Skewed multivariate models related to hidden truncation and/or selective reporting', *Test* **11**(1), 37–39.

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

Chaibub-Neto, E. & Branco, M. (2003), *Bayesian Reference Analysis for Binomial Calibration Problem*, IME-USP.

Chiogna, M. (1998), 'Some results on the scalar skew-normal distribution', *Journal Italian Statistical Society* **1**, 1–13.

DiCiccio, T. J. & Monti, A. C. (2004), 'Inferential aspects of the skew exponential power distribution', *Journal of the American Statistical Association* **99**, 439–450.

Durrans, S. R. (1992), 'Distributions of fractional order statistics in hydrology', *Water Resources Research* pp. 1649–1655.

Gupta, D. & Gupta, R. C. (2008), 'Analyzing skewed data by power normal model', *Test* **17**(1), 197–210.

Gupta, R. S. & Gupta, R. D. (2004), 'Generalized skew normal model', *Test* **13**(2), 501–524.

IDEAM (2006), *Estudio Agroclimático del Departamento de Córdoba*, Fondo Editorial Universidad de Córdoba.

Lin, G. D. & Stoyanov, J. (2009), 'The logarithmic skew-normal distributions are moment-indeterminate', *Journal of Applied Probability* **46**(3), 909–916.

Martínez-Flórez, G. (2011), Extensões do modelo α-potêncial, Tese de doutorado, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo.

R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org

Rotnitzky, A., Cox, D. R., Bottai, M. & Robins, J. (2000), 'Likelihood-based inference with singular information matrix', *Bernoulli* **6**(2), 243–284.