REVISTA DE LA FACULTAD DE INGENIERIA
www.ingenieria.uda.cl
17 (2004) 30-35

Epsilon-Skew-Normal-Generalizada

Milton A. Cortés[1], Héctor W. Gómez[1]

1. Departamento de Matemáticas, Facultad de Ingeniería, Universidad de Atacama, Chile,

E-mail: mcortes@matematica.uda.cl, hgomez@matematica.uda.cl

Resumen

En este artículo hacemos una extensión de la distribución Epsilon- Skew-Normal estudiada por Mudholkar y Hutson (2000), para esto introducimos un nuevo parámetro el cual hace más flexible la asimetría de esta distribución. Esta nueva extensión genera distribuciones unimodales y bimodales dependiendo del signo del nuevo parámetro. Estudiamos las propiedades básicas de esta nueva familia de distribuciones, a si como su representación estocástica y momentos.

Palabras claves: Asimetría, Distribución Epsilon – Skew - Normal, Kurtosis.

Abstract

In this article we make an extension of the distribution Epsilon- Skew-Normal studied by Mudholkar and Hutson (2000), for this we introduce a new parameter which makes the asymmetry more flexible of the distribution. This new extension generates unimodales and bimodales distributions depending on the sign of the new parameter. We studied the basic properties of this new family of distributions, to if like its stochastic representation and moments.

Key words: Epsilon-Skew-Normal Distribution; Kurtosis; Skewness.

1. Introducción.

Mudholkar y Hutson (2000) introducen la familia de distribuciones Epsilon-Skew-Normal $\{ESN(\varepsilon) ; |\varepsilon| < 1\}$ con parámetro de asimetría, de modo que $ESN(0)$ sea la distribución normal estándar:
Es decir $X \sim ESN(\varepsilon)$ y su densidad es:

$$f(x \mid \varepsilon) = \begin{cases} \phi\left(\dfrac{x}{(1+\varepsilon)}\right) & si \quad x < 0 \\ \phi\left(\dfrac{x}{(1-\varepsilon)}\right) & si \quad x \geq 0 \end{cases} \qquad [1]$$

Donde ϕ es la densidad normal estándar.
Algunas propiedades de esta densidad son:

$$E[X] = \frac{-4\varepsilon}{\sqrt{2\pi}} \qquad \sigma^2 = Var(X) = \frac{(3\pi - 8)\cdot\varepsilon^2 + \pi}{\pi} \qquad [2]$$

$$\sqrt{\beta_1} = \frac{2\sqrt{2\varepsilon}\cdot\left[(5\pi - 16)\cdot\varepsilon^2 - \pi\right]}{\left[(3\pi - 8)\cdot\varepsilon^2 + \pi\right]^{3/2}} \qquad [3]$$

$$\beta_2 = \frac{(15\pi^2 + 16\pi - 192)\varepsilon^2 + 10\pi(3\pi - 8)\varepsilon^2 + 3\pi^2}{\left[(3\pi - 8)\cdot\varepsilon^2 + \pi\right]^2} \qquad [4]$$

Donde $\sqrt{\beta_1}$, β_2 son los coeficientes de asimetría y kurtosis respectivamente.
De [3] y [4] se tiene que

$$-0.9953 \leq \sqrt{\beta_1} \leq 0.9953 \qquad [5]$$

$$3.0000 \leq \beta_2 \leq 3.8692$$

Recientemente, Arellano-Valle, Gómez y Quintana (2005) extienden esta familia a una clase general de distribuciones asimétricas, donde dan especial atención a la distribución Epsilon-Skew-Exponencial-Potencia.

En este artículo formulamos una extensión de la distribución Epsilon- Skew-normal, que hace más flexible la asimetría de esta densidad. Una consecuencia importante que obtenemos con esta extensión es que se construye un modelo bimodal asimétrico, es decir, se tiene una extensión del modelo generado por una mezcla de normales.

El artículo se desarrolla de la siguiente manera, en la sección 2 entregamos la definición de la distribución Epsilon-Skew-Normal-Generalizada, su función de distribución acumulada y propiedades básicas.

La representación estocástica, la extensión a una familia de localización y escala, los momentos, los coeficientes de asimetría y kurtosis son discutidos en la sección 3. El artículo finaliza con algunas conclusiones en la sección 4.

2. Extensión de la distribución Epsilon-Skew-Normal

En esta sección damos la definición de la distribución Epsilon-Skew-Normal-Generalizada y estudiamos sus propiedades.

Definición 1.

La función de densidad de probabilidad de la extensión de la epsilon-skew-normal es:

$$f(x \mid \varepsilon, \delta) = \begin{cases} \dfrac{\phi\left(\dfrac{x}{1+\varepsilon} - \delta\right)}{2(1-\Phi(\delta))} & si \quad x < 0 \\[2ex] \dfrac{\phi\left(\dfrac{x}{1-\varepsilon} + \delta\right)}{2(1-\Phi(\delta))} & si \quad x \geq 0 \end{cases} \qquad [6]$$

Donde ϕ y Φ son la densidad $N(0,1)$ y su función de distribución respectivamente con $|\varepsilon| < 1$ y $\delta \in R$.

Si X tiene una densidad como en [6] decimos que X es una variable aleatoria Epsilon-Skew-Normal-Generalizada y la denotaremos por $X \sim ESNG(\varepsilon, \delta)$.

En las Figuras 1 y 2 se muestra la forma que toma [6] en el caso unimodal y las Figuras 3, 4, 5 y 6 en el caso bimodal para diferentes valores de ε y δ .

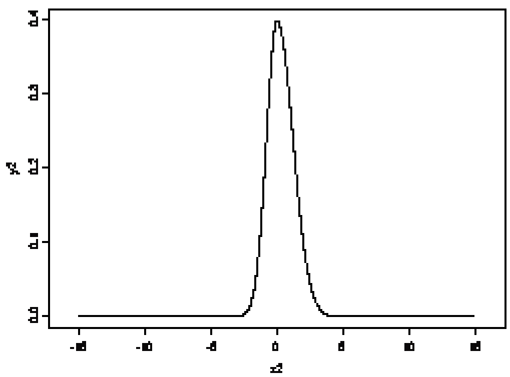

Figura 1: Función de densidad unimodal $\varepsilon = -0.2 \ y \ \delta = 0$

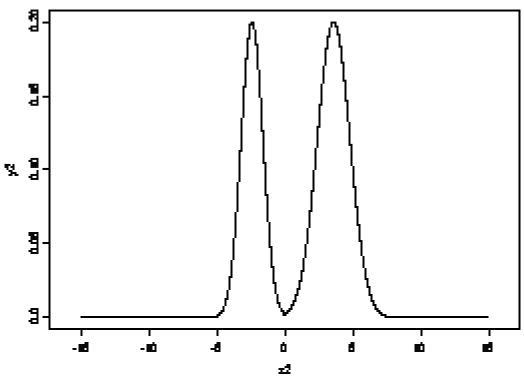

Figura 2: Función de densidad unimodal $\varepsilon = 0.2 \ y \ \delta = -3$

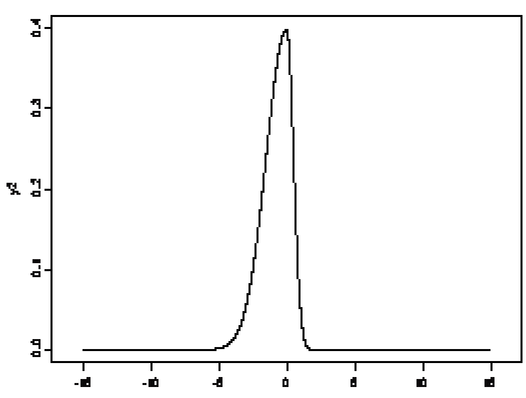

Figura 3: Función de densidad bimoda $\varepsilon = -0.2 \ y \ \delta = -3$

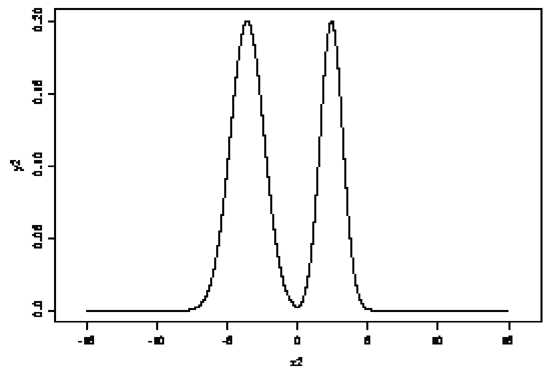

Figura 4: Función de densidad bimodal $\varepsilon = 0.5 \ y \ \delta = 0$

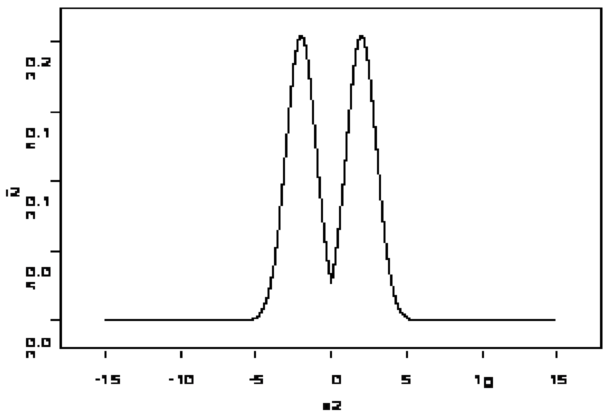

Figura 5: Función de densidad bimodal $\varepsilon = 0 \ y \ \delta = -2$

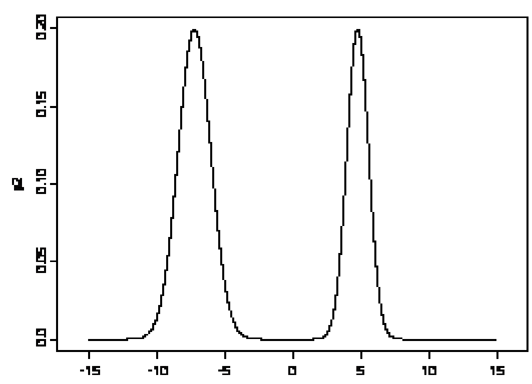

Figura 6: Función de densidad bimodal $\varepsilon = 0.2 \ y \ \delta = -6$

2.1 Función de Distribución Acumulada de la Epsilon- Skew-Normal-Generalizada.

Si X tiene densidad como en [6] entonces su función de distribución acumulada es:

$$F(x|\,\varepsilon,\delta) = \begin{cases} \dfrac{(1+\varepsilon)}{2(1-\Phi(\delta))}\phi\left(\dfrac{x}{1+\varepsilon}-\delta\right) & si \quad x<0 \\ \varepsilon + \dfrac{(1+\varepsilon)}{2(1-\Phi(\delta))}\phi\left(\dfrac{x}{1-\varepsilon}+\delta\right) & si \quad x\geq 0 \end{cases} \qquad [7]$$

Y su función de Cuantile en la forma canónica es

$$Q_0 = \begin{cases} (1+\varepsilon)\left[\delta + \Phi^{-1}\left(\dfrac{2(1-\Phi(\delta))}{(1+\varepsilon)}\cdot u\right)\right] \\ \quad si \quad 0 < u < \dfrac{1+\varepsilon}{4(1-\Phi(\delta))} \\ (1-\varepsilon)\left[-\delta + \Phi^{-1}\left(\dfrac{u-\varepsilon}{(1-\varepsilon)}\cdot 2(1-\Phi(\delta))\right)\right] \\ \quad si \quad \varepsilon + \dfrac{1-\varepsilon}{4(1-\Phi(\delta))} \leq u < \varepsilon + \dfrac{1-\varepsilon}{2(1-\Phi(\delta))} \end{cases} \qquad [8]$$

Donde $-1<\varepsilon<1$ y $\delta \in R$

2.2 Propiedades Básicas.

Las siguientes propiedades se obtienen en forma inmediata de la definición [6].

Proposición 1.

1.- $f(x|\,\varepsilon=0,\delta=0)=\phi(x)$

2.- $$f(x\,|\,\varepsilon,\delta=0)=\phi\left(\frac{x}{1+\varepsilon}\right)I\{x<0\} + \phi\left(\frac{x}{1-\varepsilon}\right)I\{x\geq 0\}$$

3.- $$\lim_{\varepsilon\to -1} f(x/\varepsilon,\delta) = \frac{\phi\left(\frac{x}{2}+\delta\right)}{2(1-\Phi(\delta))}I\{x\geq 0\}$$

4.- $$\lim_{\varepsilon\to 1} f(x/\varepsilon,\delta) = \frac{\phi\left(\frac{x}{2}-\delta\right)}{2(1-\Phi(\delta))}I\{x<0\}$$

5.- $$\lim_{\delta\to\pm\infty} f(x/\varepsilon,\delta)=0$$

Comentarios.

El resultado de la proposición 1 observamos que: [1] corresponde a la distribución normal estándar, [2] corresponde a la distribución Epsilon-Skew-Normal, [3] y [4] indican que estas densidades limites son una generalización de las densidades Half-Normal

3. Representación Estocástica

La proposición siguiente muestra la representación estocástica del modelo epsilon-skew-normal-generalizada. Es decir si X tiene densidad como en [6], luego X puede representarse como el producto de dos variables aleatorias independientes.

Proposición 2.

Para $|\varepsilon|<1$ y $\delta \in R$ se tiene que $X \sim ESNG(\varepsilon,\delta)$ si y solo si existen dos variables aleatorias independientes U_ε y V_δ con $V_\delta \sim \dfrac{\phi(x+\delta)}{1-\Phi(\delta)}I\{x\geq 0\}$ y $P(U_\varepsilon = 1-\varepsilon) = 1-P(U_\varepsilon = -(1+\varepsilon)) = \dfrac{1-\varepsilon}{2}$ tal que $X = U_\varepsilon \cdot V_\delta$

Demostración. Se obtiene directamente de la definición.

Ahora, extenderemos la distribución Epsilon-Skew-normal-Generalizada a una familia de localización y escala.

Definición 2.

Sea $X \sim ESNG(\varepsilon, \delta)$. La familia de distribuciones de localización y escala, es definida como la distribución $Z = \mu + \sigma \cdot X$, $\forall \mu \in R$, $\sigma > 0$ y su función de densidad está dada por:

$$f\left(z \middle| \boldsymbol{\theta}\right) = \begin{cases} \dfrac{1}{\sigma} \dfrac{\phi\left(\dfrac{z-\mu}{\sigma\,(1+\varepsilon)} - \delta\right)}{2(1-\Phi(\delta))} & si \quad z < \mu \\ \dfrac{1}{\sigma} \dfrac{\phi\left(\dfrac{z-\mu}{\sigma\,(1-\varepsilon)} + \delta\right)}{2(1-\Phi(\delta))} & si \quad z \geq \mu \end{cases} \qquad [9]$$

Donde $\boldsymbol{\theta} = (\mu, \sigma, \varepsilon, \delta)$ y la denotamos por: $Z \sim ESNG(\boldsymbol{\theta})$

3.1 Momentos

Proposición 3.

Sea $X \sim ESNG(0,1,\varepsilon,\delta)$ **y** $Z \sim ESNG(\mu,\sigma,\varepsilon,\delta)$**.** Entonces

$$\mu_r = \left[\frac{\left[(1-\varepsilon)^{r+1}\right] + (-1)^r (1+\varepsilon)^{r+1}}{2} \right] \frac{\sum_{k=0}^{r} \binom{r}{k} (-\delta)^{r-k}}{(1-\Phi(\delta))} M(k,\delta)$$

$$\forall\, r = 1, 2, \ldots\ldots \qquad [10]$$

Donde $\mu_r = E\left[X^r\right]$, $M(k,\delta) = \int_{\delta}^{\infty} u^k \phi(u)\,du$

y $E\left[Z^r\right] = \sum_{k=0}^{r} \binom{r}{k} \sigma^r u^{r-k} u_k$

Demostración.

El resultado se obtiene usando la independencia de U_ε y V_δ en la proposición 2 y el teorema del binomio.

Comentario.

En la Tabla 1, se muestran algunas expresiones de M para distintos valores de k y observamos que cuando k es par las expresiones dependen de funciones númericas.

k	$M(k,\delta)$
0	$1-\Phi(\delta)$
1	$\frac{1}{2}\frac{e^{-\frac{1}{2}\delta^2}}{\sqrt{\pi}}\sqrt{2}$
2	$\frac{1}{2}\frac{\sqrt{\pi} + \sqrt{2}\cdot\delta\cdot e^{-\frac{1}{2}\delta^2} - \sqrt{\pi}\cdot erf\left(\frac{1}{2}\sqrt{2}\cdot\delta\right)}{\sqrt{\pi}}$
3	$\frac{1}{2}\frac{\sqrt{2}\cdot e^{-\frac{1}{2}\delta^2}(\delta^2+2)}{\sqrt{\pi}}$
4	$\frac{1}{2}\frac{3\sqrt{\pi} + \sqrt{2}\cdot\delta^3\cdot e^{-\frac{1}{2}\delta^2} + 3\sqrt{2}\cdot\delta\cdot e^{-\frac{1}{2}\delta^2} - 3\sqrt{\pi}\cdot erf\left(\frac{1}{2}\sqrt{2}\cdot\delta\right)}{\sqrt{\pi}}$

Tabla 1: Expresiones de M para distintos valores de k

Una consecuencia inmediata de la proposición 3, son la esperanza y la varianza de la distribución $ESNG(\mu,\sigma,\varepsilon,\delta)$ dada por:

$$E[Z] = \sigma \cdot \mu + 2\sigma\varepsilon \cdot \left[\delta - \frac{1}{\sqrt{2\pi}} \frac{\exp\left(\frac{-\delta}{2}\right)}{(1-\Phi(\delta))}\right] \qquad [11]$$

$$\begin{aligned} V[Z] = \sigma^2 \Bigg[& 4\mu^2 \mu\varepsilon \left(\delta - \frac{M(1,\delta)}{(1-\Phi(\delta))}\right) \\ & + \left(1+3\varepsilon^2\right)\left(\delta^2 - \delta\frac{M(1,\delta)}{(1-\Phi(\delta))} + \frac{M(2,\delta)}{(1-\Phi(\delta))}\right) \\ & - \left(\mu + 2\varepsilon\left[\delta - \frac{M(1,\delta)}{(1-\Phi(\delta))}\right]\right)^2 \Bigg] \end{aligned} \qquad [12]$$

Por otro lado, los coeficientes de asimetría y kurtosis están dados por:

$$\sqrt{\beta_1} = \frac{\mu_3 - 3\mu_1 \cdot \mu_2 + 2\mu_1^3}{\left[(1+3\varepsilon^2)\left(\delta^2 - \delta\frac{M(1,\delta)}{(1-\Phi(\delta))} + \frac{M(2,\delta)}{(1-\Phi(\delta))}\right) - 4\varepsilon^2\left(\delta - \frac{M(1,\delta)}{(1-\Phi(\delta))}\right)^2\right]^{3/2}} \qquad [13]$$

Y

$$\beta_2 = \frac{\mu_4 - 4\mu_1 \cdot \mu_3 + 6\mu_1^2 \mu_2 - 3\mu_1^4}{\left[(1+3\varepsilon^2)\left(\delta^2 - \delta \frac{M(1,\delta)}{(1-\Phi(\delta))} + \frac{M(2,\delta)}{(1-\Phi(\delta))} \right) - 4\varepsilon^2 \left(\delta - \frac{M(1,\delta)}{(1-\Phi(\delta))} \right)^2 \right]^2} \quad [14]$$

Con μ_r, $r = 1, 2, \ldots$ son definidos en la Proposición 3

δ	**Asimetría**	**Kurtosis**
-4.0	-0.0020< $\sqrt{\beta_1}$ <0.0020	1.2283< β_2 < 2.9930
-3.0	-0.0364 < $\sqrt{\beta_1}$ <0.364	1.3782< β_2 < 2.9167
-2.0	-0.2209< $\sqrt{\beta_1}$ <0.2209	1.2283< β_2 < 2.7609
-1.0	-0.5918< $\sqrt{\beta_1}$ <0.5918	3.0575< β_2 < 3.7816
-0.8	-0.6753< $\sqrt{\beta_1}$ <0.6753	2.2407< β_2 < 3.0013
-0.5	-0.7994< $\sqrt{\beta_1}$ <0.7994	2.6059< β_2 < 3.3730
0.0	-0.9953< $\sqrt{\beta_1}$ <0.9953	3.0000< β_2 < 3.8692
0.5	-1.1689< $\sqrt{\beta_1}$ <1.1689	3.3946< β_2 < 4.4291
0.8	-1.2605< $\sqrt{\beta_1}$ <1.2605	3.6213< β_2 < 4.7719
1.0	-1.3162< $\sqrt{\beta_1}$ <1.3162	3.7659< β_2 < 4.9974
2.0	-1.5368< $\sqrt{\beta_1}$ <1.5368	4.3878< β_2 < 6.0186
3.0	-1.6792< $\sqrt{\beta_1}$ <1.6792	4.8367< β_2 < 6.8031
4.0	-1.7826< $\sqrt{\beta_1}$ < 7826	5.1320< β_2 < 2607

Tabla 2: Intervalos de asimetría y kurtosis, para distintos valores de δ.

4. Conclusiones

Podemos decir que se ha obtenido un modelo unimodal con asimetría más flexible que el modelo de Mudholkar y Hutson(2000) y un modelo bimodal asimétrico que es una extensión del modelo generado por una mezcla de normales. Este modelo bimodal nos servirá para modelar conjuntos de datos bimodales con una cierta asimetría en una de las modas.

5. Referencias

Revistas

Abramowitz, M., Stegun, I. (1972)- Handbook of Mathematical Functions-Dover Publications.

Arellano-Valle, R.B., Gómez, H.W. and Quintana, F.A.(2005). Statistical Inference for a General Class of Asymmetric Distributions. To apper in *Journal of Statistical Planning and Inference.*

Henze, N. (1986) – A Probabilistic Representation of the Skew- Normal Distribution – *Scand. J. Statist.* **13**, 271-5

Mudholkar, G.S.and Hutson, A:D.(2000) – The Epsilon- Skew-Normal Distribution for Analyzing Near- Normal Data – *Journal of Statistical Planning and Inference*. **83**, 291 -309