

UN PROBLEMA EN LA ESTIMACIÓN DEL PARÁMETRO DE FORMA DEL MODELO NORMAL - ASIMÉTRICO

Dr.(c) Omar Chocotea Poca¹ & Mgtr. Iván Yony Aliaga Casceres²

✉ omar.chocotea@postgrado.uv.cl

RESUMEN

El modelo normal-asimétrico es adecuado cuando la estructura de los datos presenta una moda y asimetría. En esta nota se observa la existencia de un problema de estimación del parámetro de forma utilizando el método de momentos en tamaños de muestra pequeños contemplando algunos aspectos de *Gupta & Gupta (2008)* [Test,17, 197–210].

PALABRAS CLAVE

Modelo asimétrico; problema de estimación

ABSTRACT

The normal-asymmetric model is suitable when the data structure presents a mode and asymmetry. This note shows the existence of an estimation problem of the shape parameter using the method of moments in small sample sizes, contemplating some aspects of *Gupta & Gupta (2008)* [Test,17, 197–210].

KEYWORDS

Asymmetric model; estimation problem

1. INTRODUCCIÓN

El interés por la clase de distribuciones normalasimétricay relacionadas ha crecido enormemente, el estudio de sus propiedades continúa, varias de sus propiedades permiten tener por ejemplo las propiedades de la distribución normal (ver *Azzalini, 2014; Arellano-Valle et al., 2018*).

La distribución normal-asimétrica de *Azzalini (1985)* contempla un parámetro de asimetría/forma: Una variable aleatoria (v.a.) Z tiene una distribución normal-asimétrica estándar con parámetro de forma $\lambda \in \mathbb{R}$, y representaremos por $Z \sim NA(\lambda)$, si su función de densidad de probabilidad (fdp) está dada por

$$\varphi_{\lambda}(z) = 2\varphi(z)\Phi(\lambda z), \quad (1)$$

donde $\varphi(\cdot)$ y $\Phi(\cdot)$ denotan la fdp y la función de distribución acumulada de la distribución normal estándar ($N(0, 1)$). Al asumir $\lambda = 0$ se tiene como caso especial a la distribución $N(0,1)$. La fdp de la distribución $NA(\lambda)$ es una densidad unimodal que está sesgada a la izquierda si $\lambda < 0$, y sesgada a la derecha si $\lambda > 0$. La esperanza y la varianza de $Z \sim NA(\lambda)$, están dadas por

$$E[Z] = \sqrt{2/\pi} \delta$$

y

$$Var[Z] = 1 - 2\delta^2/\pi,$$

¹ Instituto de Estadística, Universidad de Valparaíso, Chile y Carrera de Estadística, Universidad Mayor de San Andrés, Bolivia

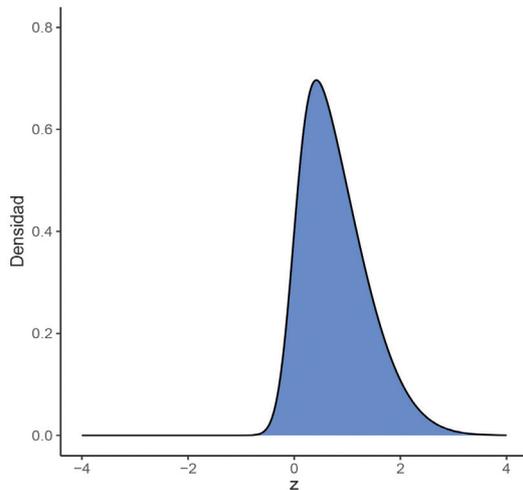
² Carrera de Estadística, Universidad Mayor de San Andrés, Bolivia

Un problema en la estimación del parámetro de forma del modelo normal-asimétrico

respectivamente, donde

$$\delta := \delta(\lambda) = \lambda/\sqrt{1 + \lambda^2} \in (-1,1)$$

Figura 1
Fdp de la $NA(\lambda)$ para $\lambda = 4$.



Fuente:Elaboración Propia

Si $Z \sim NA(\lambda)$, entonces la v.a. $Y = \xi + \omega Z$ tiene una distribución normal-asimétrica con parámetros de localización $\xi \in \mathbb{R}$, escala $\omega \in \mathbb{R}_+$, y forma $\lambda \in \mathbb{R}$, y representaremos por $Y \sim NA(\xi, \omega, \lambda)$. La fdp de $Y \sim NA(\xi, \omega, \lambda)$ está dada por

$$\varphi_{\xi, \omega, \lambda}(y) = \frac{1}{\omega} \varphi_{\lambda}((y - \xi)/\omega) \quad (2)$$

El siguiente resultado proporciona algunas propiedades para la distribución $Z \sim NA(\lambda)$.

Si $Z \sim NA(\lambda)$, entonces las siguientes propiedades son verdaderas: $\varphi_{\lambda}(0) = \varphi(0)$ para todo λ ; $-Z \sim NA(-\lambda)$, equivalente a $\varphi_{\lambda}(-z) = \varphi_{-\lambda}(z)$ para todo z ; y $Z^2 \sim \chi^2_{(1)}$, independientemente de λ .

En la Sección 2, analizamos al estimador de momentos del parámetro de asimetría de la distribución normal-asimétrica, también, en el escenario de una muestra aleatoria extraída de una población normal-asimétrica presentamos la densidad y la

función generatriz de momentos (fgm) de la media muestral. En la Sección 3, se efectúa un estudio de simulación para verificar el problema de estimación del parámetro de asimetría vía el método de momento. La Sección 4 ofrece una conclusión.

2. EL ESTIMADOR DE MOMENTO

Cuando $Z \sim NA(\lambda)$, el estimador de momento de λ es la solución de

$$\sqrt{\frac{2}{\pi}} \delta = \bar{z} \quad (3)$$

donde

$$\bar{z} = (1/n) \sum_{i \leq n} Z_i$$

es la media muestral, y la solución existe si y solamente si

$$|\bar{z}| < \sqrt{2/\pi}$$

Chen et al. (2004) incorporan las siguientes proposiciones.

Proposición 1. Sea Z_1, Z_2, \dots, Z_n una muestra aleatoria procedente de la $NA(\lambda)$,

entonces la fdp de $\bar{z} = (1/n) \sum_{i \leq n} Z_i$ está dada por

$$f(\bar{z}) = 2^n \sqrt{n} \varphi(\sqrt{n} \bar{z}) \times \Phi_n \left(\lambda \bar{z} \left[\frac{1}{1 + \lambda^2} \mathbf{I}_n + \frac{1}{n(1 + \lambda^2)} \mathbf{1} \mathbf{1}^T \right]^{1/2} \mathbf{1} \right),$$

donde $\Phi_n(\cdot)$ denota la fda de la $N_n(\mathbf{0}, \mathbf{I}_n)$, \mathbf{I}_n es la matriz identidad, y $\mathbf{1} = (1, \dots, 1)^T$.

Proposición 2. Sea Z_1, \dots, Z_n una muestra aleatoria procedente de la $NA(\lambda)$, entonces la fgm de \bar{z} está dada por

$$M_{\bar{z}}[t] = 2^n \exp\left(\frac{t^2}{2n}\right) \left[\Phi\left(\delta \frac{t}{n}\right) \right]^n$$

De las Proposiciones 1 y 2 se observa que la distribución de \bar{Z} no es conocida y no es fácil de establecer (ver Gupta & Chen, 2003; Chen et al., 2004).

Por el Teorema del Límite Central, la distribución asintótica de

$$\sqrt{n} \frac{\bar{Z} - \sqrt{2/\pi} [\lambda/\sqrt{1+\lambda^2}]}{\sqrt{1 - (2/\pi)[\lambda^2/(1+\lambda^2)]}} \quad (4)$$

es $N(0, 1)$. Por lo tanto, cuando el tamaño de la muestra n es lo suficientemente grande, aunque la distribución exacta de la media muestral estandarizada no es una distribución normal-asimétrica, su distribución asintótica sigue una distribución $N(0, 1)$ (ver Gupta & Chen, 2003; Chen et al., 2004).

3. ESTUDIO DE SIMULACIÓN

Basado en $J=1000, 5000, 10000$ conjuntos de datos, con los valores $n=10, 20, 30, 40, 50, 100, 200$ y $\lambda=1, 2, 3, 4$, se estima

$$q = Pr[|\bar{Z}| > \sqrt{2/\pi}].$$

Los resultados se muestran en el cuadro siguiente. Las figuras siguientes muestran el historial de las medias muestrales para $J = 10000$ conjuntos de datos y $n = 10, 200$, donde la banda ploma en el eje vertical inicia en $-\sqrt{2/\pi}$ y termina en $\sqrt{2/\pi}$

Los resultados indican que: a) si aumenta el tamaño de la muestra disminuye \hat{q} , y b) si aumenta parámetro de asimetría aumenta \hat{q} .

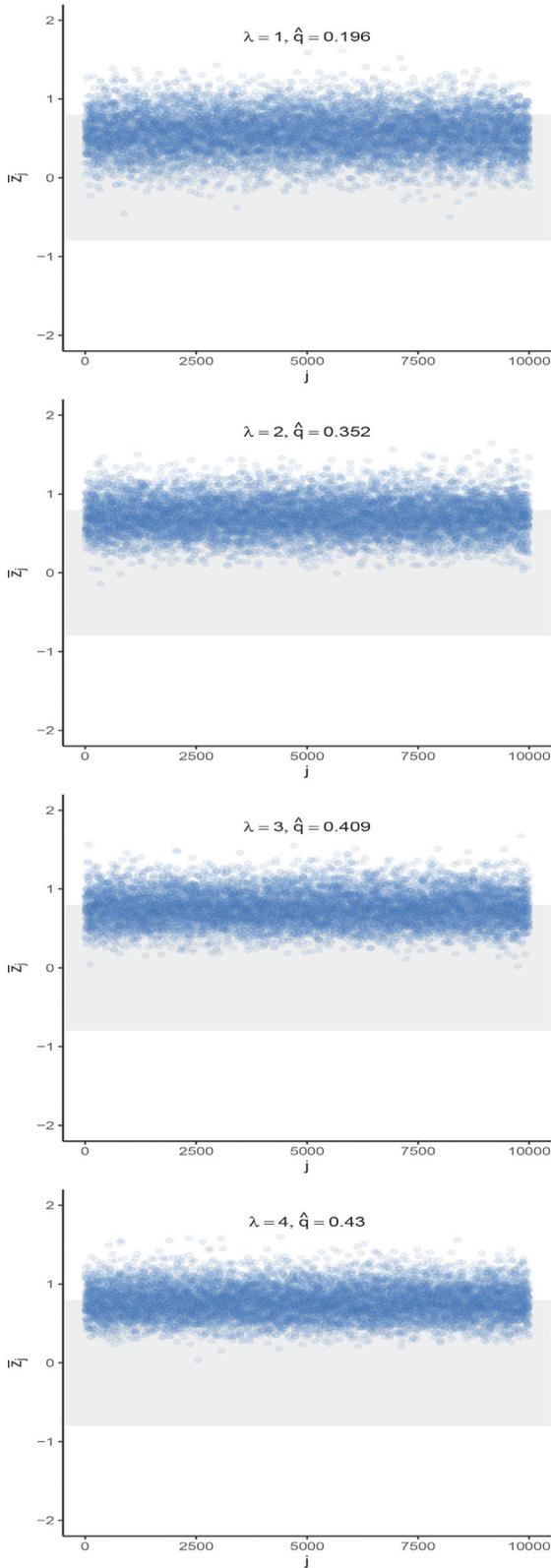
Tabla 1.
Estimación de $Pr[|\bar{Z}| > \sqrt{2/\pi}]$

J	n \ λ	1	2	3	4
1000	10	0,196	0,352	0,409	0,430
	20	0,098	0,279	0,385	0,440
	30	0,056	0,264	0,385	0,419
	40	0,044	0,229	0,303	0,379
	50	0,016	0,206	0,339	0,379
	100	0,001	0,108	0,278	0,352
	200	0,000	0,047	0,190	0,344
5000	10	0,184	0,350	0,417	0,435
	20	0,097	0,287	0,374	0,447
	30	0,059	0,250	0,356	0,404
	40	0,038	0,209	0,346	0,391
	50	0,019	0,194	0,335	0,395
	100	0,002	0,117	0,270	0,345
	200	0,000	0,047	0,194	0,347
10000	10	0,182	0,349	0,404	0,440
	20	0,100	0,290	0,387	0,416
	30	0,061	0,255	0,355	0,404
	40	0,040	0,222	0,345	0,398
	50	0,022	0,196	0,323	0,389
	100	0,002	0,117	0,260	0,348
	200	0,000	0,049	0,192	0,350

Fuente:Elaboración Propia

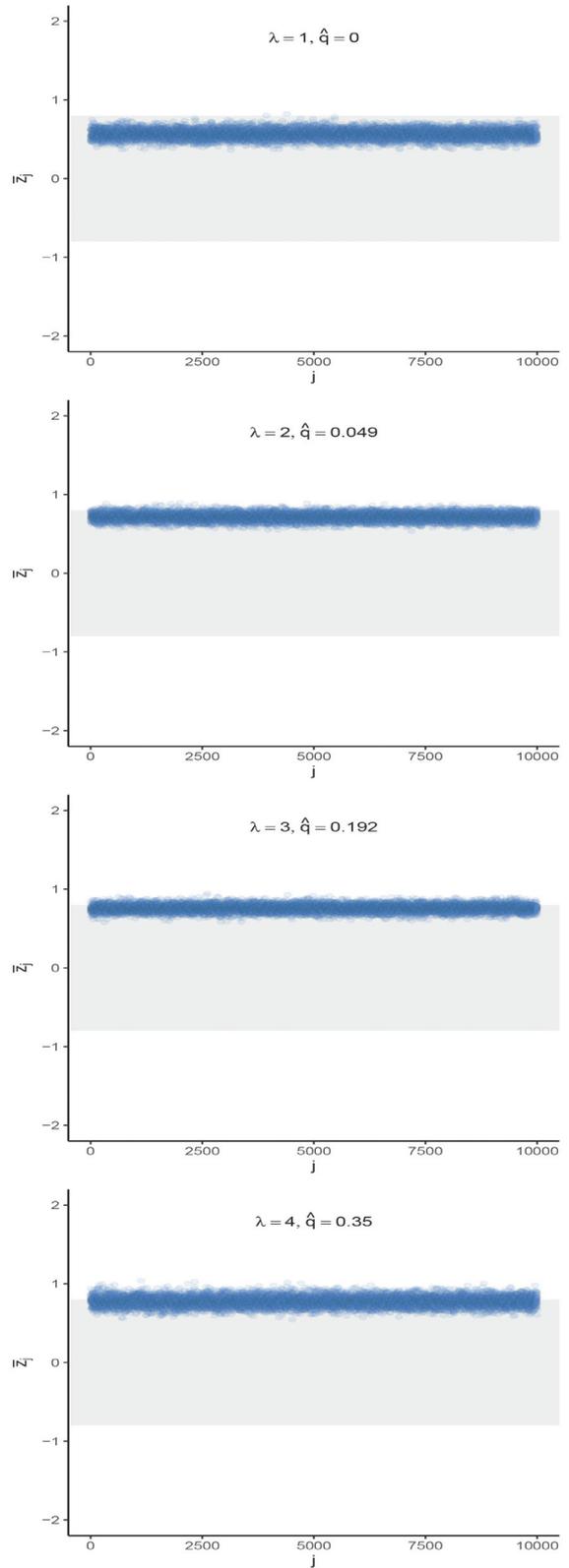
Un problema en la estimación del parámetro de forma del modelo normal-asimétrico

Figura 2.
Historial de las medias muestrales para
 $J = 10000$ y $n = 10$.



Fuente: Elaboración Propia

Figura 3.
Historial de las medias muestrales para
 $J = 10000$ y $n = 200$.



Fuente: Elaboración Propia

4. CONCLUSIÓN

En esta nota hemos analizado una opción de estimación del parámetro de asimetría de la distribución normal-asimétrica de *Azzalini* (1985), se observa que se tiene algunos problemas al estimar el parámetro de asimetría en tamaños de muestra moderados. Otros problemas de estimación pueden revisarse en *Pewsey* (2000), *Monti* (2003), *Liseo & Loperfido* (2006), y *Sartori* (2006).

Algunas alternativas de estimación se ven por ejemplo en *Azzalini* (2014) y *Arellano-Valle et al.* (2018).

AGRADECIMIENTOS

El primer autor fue parcialmente apoyado por la beca FIB-UV de la Universidad de Valparaíso, Chile. Los autores agradecen al editor por los útiles comentarios

BIBLIOGRAFÍA

1. Arellano-Valle, R. B., Ferreira, C. S., & Genton, M. G. (2018). Scale and Shape Mixtures of Multivariate Skew-Normal Distributions. *Journal of Multivariate Analysis*, 166, 98–110. 1, 4
2. Azzalini, A. (1985). A Class of Distributions which Includes The Normal Ones. *Scandinavian Journal of Statistics*, 12(2), 171–178. 1, 4
3. Azzalini, A. (2014). *The Skew-Normal and Related Families*. Cambridge: Cambridge University Press. 1, 4
4. Chen, J. T., Gupta, A. K., & Nguyen, T. T. (2004). The Density of The Skew Normal Sample Mean and its Applications. *Journal of Statistical Computation and Simulation*, 74(7), 487–494. 2
5. Gupta, A. & Chen, T. (2003). On The Sample Characterization Criterion for Normal Distributions. *Journal of Statistical Computation and Simulation*, 73(3), 155–163. 2
6. Gupta, R. D. & Gupta, R. C. (2008). Analyzing Skewed Data by Power Normal Model. *Test*, 17, 197–210. 1
7. Liseo, B. & Loperfido, N. (2006). A Note on Reference Priors for The Scalar Skew-Normal Distribution. *Journal of Statistical Planning and Inference*, 136(2), 373–389. 4
8. Monti, A. C. (2003). A Note on The Estimation of The Skew Normal and The Skew Exponential Power Distributions. *METRÓN - International Journal of Statistics*, LXI(2), 205–219. 4
9. Pewsey, A. (2000). Problems of Inference for Azzalini 's Skewnormal Distribution. *Journal of Applied Statistics*, 27(7), 859–870. 4
10. Sartori, N. (2006). Bias Prevention of Maximum Likelihood estimates for Scalar Skew Normal and Skew t Distributions. *Journal of Statistical Planning and Inference*, 136(12), 4259–4275. 4