SORT 32 (1) January-June 2008, 57-66

A note on the likelihood and moments of the skew-normal distribution

Eliseo H. Martínez[1], Héctor Varela[2],
Héctor W. Gómez[3] and Heleno Bolfarine

Abstract

In this paper an alternative approach to the one in Henze (1986) is proposed for deriving the odd moments of the skew-normal distribution considered in Azzalini (1985). The approach is based on a Pascal type triangle, which seems to greatly simplify moments computation. Moreover, it is shown that the likelihood equation for estimating the asymmetry parameter in such model is generated as orthogonal functions to the sample vector. As a consequence, conditions for a unique solution of the likelihood equation are established, which seem to hold in more general setting.

MSC: 62F10, 65C20

Keywords: Skew-normal distribution, maximum likelihood equations, moments

1 Introduction

The density function of the skew-normal distribution with parameters ε, $\omega > 0$ and λ (location, scale and asymmetry parameters, respectively) introduced by Azzalini (1985), is given by

[1] Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Chile. E-mail: emartinez@uantof.cl. Partially supported by Grant DIRINV 1323-06.

[2] Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Chile. E-mail: hvarela@uantof.cl. Partially supported by Grant DIRINV 1323-06.

[3] Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Chile. E-mail: hgomez@mat.uda.cl. Partially supported by Grant FONDECYT 1060727.

[4] Departamento de Estatistica, IME, Universidad de Sao Paulo, Brasil. E-mail: hbolfar@ime.usp.br. Partially supported by CNPq-Brasil.

Received: December 2006
Accepted: November 2007

$$\phi(x;\varepsilon,\omega,\lambda) = \frac{2}{\omega}\phi\left(\frac{x-\varepsilon}{\omega}\right)\Phi\left(\frac{x-\varepsilon}{\omega}\cdot\lambda\right),$$

where ϕ and Φ are, respectively, the density and cumulative distribution functions of the standard normal distribution. We use the notation $X \sim SN(\varepsilon,\omega,\lambda)$ to denote this distribution, and in the special case of the standard version, $X \sim SN(0,1,\lambda)$. As is well known, the density function of the ordinary normal distribution follows as a special case (i.e., $\lambda = 0$). Moreover, the k-th moment of a random variable X with $SN(\varepsilon,\omega,\lambda)$ distribution can be computed directly from the integral

$$\mathbf{E}\left[X^k\right] = \int_{-\infty}^{\infty} x^k \frac{2}{\omega}\phi\left(\frac{x-\varepsilon}{\omega}\right)\Phi\left(\frac{x-\varepsilon}{\omega}\cdot\lambda\right) dx,$$

which is quite complicated to deal with. Using this general expectation, moments can be computed as, for example, if $X \sim SN(0,1,\lambda)$, then

$$\mu_1 = \mathbf{E}\left[X\right] = \frac{\sqrt{2}\,\lambda}{\sqrt{\pi}\,\sqrt{(\lambda^2+1)}}\ ;\ \mu_2 = \mathbf{E}\left[X^2\right] = 1.$$

It is also well known that if $X \sim SN(0,1,\lambda)$ then $X^2 \sim \chi^2_{(1)}$ (Azzalini, 1985), so that the even moments are equal to the moments of the ordinary normal distribution. Henze (1986) employs a stochastic representation for this model, so that if $X \sim SN(0,1,\lambda)$, then

$$X = \delta|Z_0| + (1-\delta^2)^{1/2}Z_1,$$

where $\delta = \lambda/(1+\lambda^2)^{1/2}$ and Z_0 and Z_1 are independent $N(0,1)$ random variates. Odd moments can then be computed in a more straightforward way.

Azzalini and Capitanio (1999) report on some applications of the multivariate skew-normal distribution while Genton et al. (2001) derive the moments for the multivariate version. Pewsey (2000) discusses inference problems for the univariate skew-normal distribution such as unbounded likelihood and singular information matrix. A recent review encompassing most of the recent advances on the topic is found in Azzalini (2005).

The main object of this note is to present alternative approaches for deriving the odd moments of the skew-normal distribution and study the behaviour of the likelihood equation for the asymmetry parameter. In particular, conditions for a unique finite root are discussed. We propose using a Pascal type triangle, which seems to greatly simplify computing the odd moments. This approach seems to be new and completely overlooked by the literature.

The paper is organized as follows. Section 2 deals with odd moments of the skew-normal distribution, which are derived using a Pascal type triangle, different from the approach considered in Henze (1986). Section 3 is devoted to the study of the maximum likelihood estimator for the $SN(0, 1, \lambda)$ distribution. In Section 4 final comments are presented.

2 Odd moments of the skew-normal distributions

In this section we discuss alternative derivations for the odd moments of the skew-normal distribution. We recall that the even moments are the same as the ones for the symmetric normal model. We start by presenting a recursive relation for moments computation.

Proposition 1 *Let $X \sim SN(0, 1, \lambda)$. Then*

$$\mu_{2n+1} = \mathbf{E}\left[X^{2n+1}\right] = 2n\mu_{2n-1} + \sqrt{\frac{2}{\pi}}\frac{(2n)!}{2^n(n)!}\frac{\lambda}{(1+\lambda^2)^{n+1/2}}, \tag{1}$$

for $n = 1, 2, \ldots$.

Proof. Solving $\mu_{2n-1} = \int_{-\infty}^{\infty} x^{2n-1} 2\phi(x)\Phi(\lambda x)\, dx$ using integration by parts, we obtain

$$\mu_{2n-1} = \left[\frac{x^{2n}}{n}\phi(x)\Phi(\lambda x)\right]_{-\infty}^{\infty} + \frac{\mu_{2n+1}}{2n} - \frac{1}{2n}\int_{-\infty}^{\infty} 2\lambda\phi(x)\phi(\lambda x)\, x^{2n}\, dx.$$

Solving for the integral on the right hand side, we obtain

$$\mu_{2n-1} = \frac{\mu_{2n+1}}{2n} - \frac{1}{2n}\sqrt{\frac{2}{\pi}}\frac{(2n)!}{2^n(n)!}\frac{\lambda}{(1+\lambda^2)^{n+1/2}} \tag{2}$$

and the result follows by re-expressing the moment $\mu_{2n+1} = \mu_{2n+1}$ in the last expression. □

Equation (1) allows computing moments for the skew normal distribution in a recursive fashion, which can be troublesome. We can, however, use a more direct approach, as described next. Recall the double factorial function for a positive integer:

$$n!! = \begin{cases} n(n-2)\cdots 3\cdot 1 & n > 0, \textit{ odd} \\ n(n-2)\cdots 4\cdot 2 & n > 0, \textit{ even} \\ 1 & n = 0, 1. \end{cases}$$

Of its many properties, the one we may use is

$$(2n-1)!! = \frac{(2n)!}{2^n n!},$$

so that recursive equation (2) can be written as

$$\mu_{2n+1} = 2n\mu_{2n-1} + \sqrt{\frac{2}{\pi}}(2n-1)!!\frac{\lambda}{(1+\lambda^2)^{n+1/2}},$$

for $n = 1, 2, \ldots$.

The first few iterations of this equation lead to the following moments:

$$\mu_1 = \frac{\sqrt{2}\,\lambda}{\sqrt{\pi}\ \sqrt{(\lambda^2+1)}},$$

$$\mu_3 = \sqrt{2/\pi}\left[2\frac{\lambda}{(1+\lambda^2)^{1/2}} + 1!!\frac{\lambda}{(1+\lambda^2)^{3/2}}\right],$$

$$\mu_5 = \sqrt{2/\pi}\left[2\cdot 4\frac{\lambda}{(1+\lambda^2)^{1/2}} + 4\cdot 1!!\frac{\lambda}{(1+\lambda^2)^{3/2}} + 3!!\frac{\lambda}{(1+\lambda^2)^{5/2}},\right]$$

and

$$\mu_7 = \sqrt{2/\pi}\left[2\cdot 4\cdot 6\frac{\lambda}{(1+\lambda^2)^{1/2}} + 6\cdot 4\cdot 1!!\frac{\lambda}{(1+\lambda^2)^{3/2}} + 6\cdot 3!!\frac{\lambda}{(1+\lambda^2)^{5/2}} + 5!!\frac{\lambda}{(1+\lambda^2)^{7/2}}\right],$$

leading to the following compact equation:

$$\mu_{2n+1} = \sqrt{\frac{2}{\pi}}\sum_{k=0}^{n}\frac{(2n)!!}{(2k)!!}\frac{(2k-1)!!\lambda}{(1+\lambda^2)^{(2k+1)/2}},$$

which can be rewritten after a straightforward factorization as

$$\mu_{2n+1} = \sqrt{2/\pi}\frac{\lambda}{(1+\lambda^2)^{(2n+1)/2}}\sum_{k=0}^{n}\frac{(2n)!!}{(2k)!!}(2k-1)!!(1+\lambda^2)^{n-k}.$$

Expanding $(1+\lambda^2)^{n-k}$ leads to

$$\mu_{2n+1} = \sqrt{2/\pi}\frac{\lambda}{(1+\lambda^2)^{(2n+1)/2}}\sum_{k=0}^{n}\sum_{j=0}^{n-k}\frac{(2n)!!}{(2k)!!}(2k-1)!!\binom{n-k}{j}\lambda^{2j}.$$

It can be shown after extensive but straightforward algebraic manipulations that the above expressions and the expression given in Hence (1986), Corollary 4, are equivalent. Hence, using this last expression, the first four odd moments of the skew normal distribution can be written as

$$\mu_1 = \sqrt{2/\pi}\frac{1}{(1+\lambda^2)^{1/2}}\lambda \cdot 1!!$$

$$\mu_3 = \sqrt{2/\pi}\frac{1}{(1+\lambda^2)^{3/2}}(2!!\lambda^3 + 3!!\lambda)$$

$$\mu_5 = \sqrt{2/\pi}\frac{1}{(1+\lambda^2)^{5/2}}(4!!\lambda^5 + 4(2!! + 3!!)\lambda^3 + 5!!\lambda)$$

$$\mu_7 = \sqrt{2/\pi}\frac{1}{(1+\lambda^2)^{7/2}}(6!!\lambda^7 + 6[4!! + 4(2!! + 3!!)]\lambda^5 + 6[4(2!! + 3!!) + 5!!]\lambda^3 + 7!!\lambda)$$

and by fixing attention to the coefficients of λ^{2j+1}, a Pascal type triangle formation can be depicted. The results are then formalized in the next proposition.

Proposition 2 *If $X \sim SN(0,1,\lambda)$ and n is a positive integer then the odd moments are given by*

$$\mathbf{E}\left[X^{2n-1}\right] = \frac{\sqrt{2}}{\sqrt{\pi}\,(\lambda^2+1)^{(2n-1)/2}}\,Q_n(\lambda)$$

where

$$Q_n(\lambda) = \sum_{k=1}^{n} a_n(k)\,\lambda^{2k-1}$$

and the coefficients $a_n(k)$ are computed iteratively by using

$$\begin{aligned} a_1(1) &= 1; \\ a_n(1) &= (2n-1)\,a_{n-1}(1)\,;\; n \geq 2; \\ a_n(k) &= (2n-2)\,(a_{n-1}(k) + a_{n-1}(k-1))\,;\; n \geq 2, 1 < k < n; \\ a_n(n) &= (2n-2)\,a_{n-1}(n-1). \end{aligned}$$

Hence, the odd moments of X can be computed iteratively by computing the coefficients $a_n(k)$ which can be cumbersome. So, it would be interesting to have a more convenient and direct way, without having to use recursion.

We show in the sequel that the coefficients $a_n(k)$ can be computed by using a Pascal type triangle, constructed according to the following steps. It starts with 1 as the ordinary Pascal triangle, and the left hand side of the triangle has even levels greater than 1 (2,4,6,...), for example, level 2 has two numbers, 2 and 3 and the right hand side of the triangle has odd levels greater than 1 (3,5,7,...). For example, the leftmost number in level 3 is number 8, which results from multiplying basis number (2) by the following even number which is 4. On the other hand, the rightmost number in that level follows by multiplying the basis number (3) by the following odd number which is 5, so that 15 results.

Moreover, all the other numbers are generated similarly as the Pascal Triangle except that each resulting number is multiplied by the next corresponding even level. For example, 20 is obtained by adding 2 and 3 (equals 5), which multiplied by 4 makes 20. 168 is obtained by adding 8 and 20 and multiplying by the next even level which is 6, leading to 168. Similarly, 210 is obtained by adding 20 and 15 and also multiplying by the line even level, 6, and so on.

The polynomial $Q_n(\lambda)$ can then be computed for odd degrees. It also presents a form of regularity which we believe has not been noticed so far in the literature. It is expressed by the fact that the odd moments of the polynomial $Q_n(\lambda)$, of which the first and last coefficients, namely $a_n(1)$ and $a_n(n)$, are odd and even double factorials, respectively, and the remaining coefficients follow a Pascal type triangle regularity and are, consequently, easily computable.

As an example, we compute in the sequel the coefficients for the moments of order 1, 3, 5, 7, 9 for the skew normal distribution. As described above, we obtain:

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 2 & & 3 & & & \\
 & & 8 & & 20 & & 15 & & \\
 & 48 & & 168 & & 210 & & 105 & \\
384 & & 1728 & & 3024 & & 2520 & & 945
\end{array}
$$

As an illustration, we compute the seventh moment of the skew-normal distribution, which follows by considering the fourth line ($n = 4$) in the triangle above, leading to

$$Q_4(\lambda) = 48\lambda^7 + 168\lambda^5 + 210\lambda^3 + 105\lambda,$$

from which the seventh moment can be computed as

$$\mathbf{E}\left[X^7\right] = \frac{\sqrt{2}\lambda\,(48\lambda^6 + 168\lambda^4 + 210\lambda^2 + 105)}{\sqrt{\pi}(\lambda^2 + 1)^{7/2}}.$$

Note that the right hand side of the triangle reproduces the even moments for the skew normal distribution.

Using the above results, moments for the location-scale situation, that is, for $Z \sim SN(\varepsilon, \omega, \lambda)$, can be easily computed by using the fact that

$$E[Z^r] = \sum_{k=0}^{r} \binom{r}{k} \varepsilon^{r-k} \omega^k \mu_k.$$

which follows directly by using the binomial expansion.

3 Maximum likelihood equation for the asymmetry parameter

Suppose that $X \sim SN(0, 1, \lambda)$ with λ unknown, and let $x_1, \cdots, x_n$ a random sample from this distribution. The likelihood function for the parameter λ based on the observed sample is then given by

$$l(\lambda) = n \log 2 + \sum_{i=1}^{n} \log \phi(x_i) + \sum_{i=1}^{n} \log \Phi(\lambda x_i)$$

Differentiating the log of the likelihood with respect to parameter λ, the following equation is obtained:

$$\frac{\partial l(\lambda)}{\partial \lambda} = \sum_{i=1}^{n} R(i) x_i = 0, \tag{3}$$

where

$$R(i) = \frac{\phi(\lambda x_i)}{\Phi(\lambda x_i)}. \tag{4}$$

The estimator of λ can be found easily. In fact, we only have to solve the equation:

$$\sum_{i=1}^{n} R(i) x_i = 0.$$

Hence, in this case, to find the maximum likelihood estimator of λ, we have to find the $R(i)$ (or λ) defined according to (4), so that the vector $\mathbf{R} = (R(1), \cdots, R(n))$ is orthogonal to $\mathbf{x} = (x_1, \cdots, x_n)$. Note that, $\mathbf{Rx}$ is a function of λ, that is, we can write

$$f(\lambda) = \mathbf{R} \cdot \mathbf{x}, \tag{5}$$

and the maximum likelihood estimator of λ makes **R** and **x** orthogonal. This function clearly is continuous for all values of λ. Conditions for a unique root in (5) are established next.

Proposition 3 *If* $\mathbf{x} = (x_1, \cdots, x_n)'$*, and if there exists k and j such that* $x_k < 0$ *and* $x_j > 0$ *then* $f(\lambda) = 0$ *has a unique real root.*

Proof. We can separate in the likelihood function the positive and negative sample elements, so that $f(\lambda) = 0$ if and only if

$$f(\lambda) = \sum_{x_i > 0} R(i)\, x_i + \sum_{x_j < 0} R(j)\, x_j = 0$$

Moreover, defining

$$h(\lambda) = \sum_{x_i > 0} R(i) x_i \;;\; g(\lambda) = \sum_{x_j < 0} R(j)\, x_j,$$

it clearly follows that $h(\lambda) > 0$ and $g(\lambda) < 0$ for any λ. Notice that $R(x) > 0$ and it is a differentiable monotonically decreasing function from $+\infty$ to 0 as x ranges from $-\infty$ to $+\infty$. Hence, $h(-\infty) = \infty$ and $h(+\infty) = 0$ so that the proof can proceed as follows. Notice that $f(\lambda)$ is the sum of a positive decreasing function h and a negative increasing function g, such that

$$f(-\infty) = \infty + 0 = \infty, \quad f(\infty) = 0 - \infty = -\infty$$

and therefore $f(\lambda)$ has a root; monotonicity of h and g imply uniqueness of this root. □

Liseo (1990) has noticed the inverse statement to Proposition 3, namely that if all sample elements have the same sign, then the MLE does not exist. Pewsey (2006) considers the general case, that is, the function $\Phi(.)$ is replaced by a general G.

We call attention to the fact that the above results can be extended to any density of the form

$$h(x) = 2g(x)\Phi(\lambda x),$$

where $g(.)$ is a symmetric density function and $\Phi(.)$ is as above. Such more general families of densities are considered in Arellano-Valle and Genton (2005) and Gómez *et al.* (2007).

4 Discussion

In this note we have shown new approaches for computing odd moments of the skew normal distribution. The first method uses a Pascal type triangle making computing the moments more accessible, while the second method uses a recursive approach. Further, it is also shown that the maximum likelihood equation for estimating the asymmetry parameter λ is a product of an orthogonal function to the sample vector, leading to conditions for a unique root of the likelihood equation. Clearly, Proposition 3 holds in more general settings.

Acknowledgments

The authors acknowledge helpful comments and suggestions from two referees which substantially improved the presentation.

References

Arellano-Valle, R. B. and Genton, M. G. (2005). On fundamental skew distributions. *Journal of Multivariate Analysis*, 96, 93-116.

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

Azzalini, A. (2005). The skew-normal distribution and related multivariate families (with discussion). *Scandinavian Journal of Statistics*, 32, 159-188.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society*, B, 61, 579-602.

Genton, M. G., He, L. and Liu. X. (2001), Moments of skew-normal random vectors and their quadratic forms. *Statististics and Probability Letters*, 51, 319-325.

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

Henze, N. (1986). A probabilistic representation of the skew-normal distribution. *Scandinavian Journal of Statistics*, 13, 271-275.

Liseo, B. (1990). La classe delle densita normali sghembe: aspetti inferenziali da un punto di vista bayesiano. *Statistica*, L, 59-70.

Pewsey, A. (2000). Problems of inference for Azzalini's skew-normal distribution. *Journal of Applied Statistics*, 27, 859-870. 13, 271-275.

Pewsey, A. (2006). Some observations on a simple means of generating skew distributions, in *Advances in Distribution Theory, Order Statistics and Inference*, N. Balakrishnan, E. Castillo and J. M. Sarabia (eds.), Boston, MA: Statistics for Industry and Technology, Birkhauser, 75-84.