A Probabilistic Representation of the 'Skew-Normal' Distribution

Author(s): Norbert Henze

Source: *Scandinavian Journal of Statistics*, Vol. 13, No. 4 (1986), pp. 271-275
Published by: Wiley on behalf of Board of the Foundation of the Scandinavian Journal of Statistics
Stable URL: https://www.jstor.org/stable/4616036
Accessed: 30-08-2018 15:46 UTC

Scand J Statist 13: 271–275, 1986

A Probabilistic Representation of the 'Skew-normal' Distribution

NORBERT HENZE
University of Hannover

ABSTRACT. In a recent paper Azzalini (1985) introduced the class $\mathcal{SN}=\{\mathcal{SN}(\lambda): \lambda\in\mathbb{R}\}$ of *skew-normal* probability distributions and studied its main properties. The salient features of the class $\mathcal{SN}$ are mathematical tractability and strict inclusion of the Normal law (for $\lambda=0$). The shape parameter λ, to some extent, controls the index of skewness. It is the purpose of this note to give a probabilistic representation of the distribution $\mathcal{SN}(\lambda)$ in terms of Normal and truncated Normal laws. This representation reveals the "structure" of the class $\mathcal{SN}$ and indicates the kind of departure from normality. The moments of a random variable Z_λ with the distribution $\mathcal{SN}(\lambda)$ are explicitly determined, and an efficient method for the Monte Carlo generation of Z_λ is shown.

Key words: normal distribution, truncated normal distribution, skewness, random number generation

1. Introduction

Let φ and Φ denote the standard Normal density and distribution function, respectively. A random variable Z_λ is said to have a *skew-normal* distribution with parameter λ, $-\infty<\lambda<\infty$, if the density of Z_λ is

$$\varphi(z;\lambda)=2\,\varphi(z)\,\Phi(\lambda z) \qquad (-\infty<z<\infty), \tag{1}$$

and we write $Z_\lambda\sim\mathcal{SN}(\lambda)$.

The class $\mathcal{SN}=\{\mathcal{SN}(\lambda): \lambda\in\mathbb{R}\}$ was introduced by Azzalini (1985), who studied its basic properties analytically, i.e. by means of the defining density (1). The salient features of the skew-normal class are mathematical tractability (the likelihood equations for a simple random sample of size n from $Y_\lambda:=\lambda_1+\lambda_2 Z_\lambda$ $(\lambda_2>0)$ are readily written down, and the Fisher information for $(\lambda, \lambda_1, \lambda_2)$ is easily computed) and "strict inclusion" of the Normal density (for $\lambda=0$).

However, the class $\mathcal{SN}$ does not meet a third "ideal" requirement, namely wide range of the indices of skewness and kurtosis. To this end, Azzalini proposed an extension of $\mathcal{SN}$ by introducing an additional shape parameter ξ; $-\infty<\xi<\infty$ (see also O'Hagan & Leonhard, 1976). This broader class of distributions, which will for brevity be termed $\mathcal{SNE}=\{\mathcal{SNE}(\lambda,\xi): \lambda, \xi\in\mathbb{R}\}$ is defined by the densities

$$\varphi(z;\lambda,\xi)=\varphi(z)\,\Phi(\lambda z+\xi)/\Phi\{\xi(1+\lambda^2)^{-1/2}\} \qquad (-\infty<z<\infty) \tag{2}$$

and we write $Z_{\lambda,\xi}\sim\mathcal{SNE}(\lambda,\xi)$, if a random variable $Z_{\lambda,\xi}$ has density (2). Observe that $\mathcal{SN}=\{\mathcal{SNE}(\lambda,0): \lambda\in\mathbb{R}\}$.

In the following we give representation theorems for the classes $\mathcal{SN}$ and $\mathcal{SNE}$ in terms of Normal and truncated Normal distributions. These representations indicate the kind of departure from the Normal law and reveal the "intrinsic structure" of $\mathcal{SN}(\lambda)$ and $\mathcal{SNE}(\lambda,\xi)$ distributions. They also provide probabilistic proofs for most of the properties A–I of the family $\mathcal{SN}$ given by Azzalini. Moreover, an explicit formula for the moments of Z_λ is obtained, and it is shown that two uniform random numbers suffice to generate Z_λ in Monte Carlo experiments.

2. Results

In what follows we write $V \sim \mathcal{N}(\mu, \sigma^2)$ if the random variable V is normally distributed with mean μ and variance σ^2. $\mathcal{T}(c)$ denotes the distribution of a standard Normal random variable truncated from below at c, $-\infty<c<\infty$, i.e. a random variable $T \sim \mathcal{T}(c)$ has density $\{1-\Phi(c)\}^{-1}\varphi(t)$ $(c \leqslant t<\infty)$.

Theorem 1

Let U, V be independent standard Normal random variables. Then

$$Z_\lambda := \frac{\lambda}{\sqrt{1+\lambda^2}}|U| + \frac{1}{\sqrt{1+\lambda^2}} V \sim \mathcal{SN}(\lambda).$$

First proof (analytical). Letting

$$a=\lambda(1+\lambda^2)^{-1/2}, \quad b=(1+\lambda^2)^{-1/2}, \tag{3}$$

we have

$$\begin{aligned} P(Z_\lambda \leqslant z) &= E[P(Z_\lambda \leqq z \mid |U|)] \\ &= \int_0^\infty P\{V \leqslant (z-au)/b\}\, 2\varphi(u)\, du \\ &= 2\int_0^\infty \Phi\{(z-au)/b\}\, \varphi(u)\, du, \end{aligned}$$

and from the relation $a^2+b^2=1$ it easily follows that

$$\begin{aligned} \frac{d}{dz} P(Z_\lambda \leqslant z) &= 2\varphi(z) \int_0^\infty (2\pi b^2)^{-1/2} \exp\{-b^{-2}(u-az)\} du \\ &= 2\varphi(z)\left\{1-\Phi\left(-\frac{a}{b}z\right)\right\} \\ &= 2\varphi(z)\,\Phi(\lambda z). \end{aligned}$$

Second proof (probabilistic). Let X and Y be independent standard Normal random variables. Then (this was the starting point in Azzalini's paper!) $\mathcal{SN}(\lambda)$ is the distribution of Y *conditionally on*

$$X<\lambda Y. \tag{4}$$

Upon performing the rotations

$$\begin{aligned} R &= (1+\lambda^2)^{-1/2}(X-\lambda Y) \\ S &= (1+\lambda^2)^{-1/2}(\lambda X+Y), \end{aligned} \tag{5}$$

R and S are independent standard Normal random variables, and (4) is equivalent to the condition $R<0$. Since

$$Y=(1+\lambda^2)^{-1/2}(-\lambda R+S) \tag{6}$$

and the conditional distribution of $-R$ given that $R<0$ equals that of $|U|$ figuring in the statement of theorem 1, we are done.

An efficient method for generating a random variable with distribution $\mathcal{SN}(\lambda)$ is given by the following corollary.

Corollary 2
Let W_1, W_2 be independent random variables, each with uniform distribution in $(0, 1)$, and let a, b be as in (3). Then

$$a\{-2\log(W_1)\}^{1/2}\sin(\pi W_2)+b(-2\log(W_1))^{1/2}\cos(\pi W_2)\sim\mathcal{SN}(\lambda).$$

Proof. Observe that the two summands are independent, with

$$(-2\log(W_1))^{1/2}\sin(\pi W_2)\sim|U|$$
$$(-2\log(W_1))^{1/2}\cos(\pi W_2)\sim V$$

(U, V figuring in the statement of theorem 1).

Theorem 1 provides probabilistic proofs for most of the properties of the family $\mathcal{SN}$ given by Azzalini (1985) (properties A, B, C, E, H and I of his paper). For example, we have the following result.

Corollary 3
Assume $Z\sim\mathcal{SN}(\lambda)$, $W\sim\mathcal{N}(0, \sigma^2)$, and let Z and W be independent. Then

$$(1+\sigma^2)^{-1/2}(Z+W)\sim\mathcal{SN}\{\lambda(1+\sigma^2+\lambda^2\sigma^2)^{-1/2}\}.$$

Proof. Let U, V_1, V_2 be independent standard Normal random variables, let a, b as in (3), and put $\tau=(1+\sigma^2)^{1/2}$. By theorem 1,

$$\frac{1}{\tau}(Z+W)\sim\frac{a}{\tau}|U|+\frac{1}{\tau}(bV_1+\sigma V_2)$$
$$=\frac{a}{\tau}|U|+\frac{1}{\tau}(b^2+\sigma^2)^{1/2}V,$$

where

$$V=(b^2+\sigma^2)^{-1/2}(bV_1+\sigma V_2)\sim\mathcal{N}(0,1),$$

and U and V are independent. Solving the equation $\eta(1+\eta^2)^{-1/2}=a/\tau$, the result follows.

Another consequence of theorem 1 is the following formula for the odd moments of a random variable $Z_\lambda\sim\mathcal{SN}(\lambda)$ (the even moments are of no interest because Z_λ^2 has the chi-square distribution χ_1^2).

Corollary 4
Let $Z_\lambda\sim\mathcal{SN}(\lambda)$. Then, for $k=0, 1, 2, \ldots$,

$$E[Z_\lambda^{2k+1}]=\sqrt{\frac{2}{\pi}}\lambda(1+\lambda^2)^{-(k+1/2)}2^{-k}(2k+1)!\sum_{\nu=0}^{k}\frac{\nu!(2\lambda)^{2\nu}}{(2\nu+1)!(k-\nu)!}.$$

Proof. By theorem 1,

$$E[Z_\lambda^{2k+1}] = E[(a|U|+bV)^{2k+1}]$$

$$= \sum_{j=0}^{2k+1} \binom{2k+1}{j} a^j b^{2k+1-j} E[|U|^j]\, E[V^{2k+1-j}].$$

Now use

$$E[V^{2r+1}]=0,$$

$$E[V^{2r}] = \frac{(2r)!}{2^r r!}$$

$$E[|U|^{2r+1}] = \sqrt{\frac{2}{\pi}}\, r!2^r \qquad \text{for } r=0, 1, 2, \ldots$$

To obtain a representation for the class $\mathscr{SNE}$ we first prove

Lemma 5
Assume X and Y are independent standard Normal random variables, and let λ, $\xi \in \mathbb{R}$. Then the distribution of Y, conditionally on

$$X < \lambda Y + \xi \tag{7}$$

is $\mathscr{SNE}(\lambda, \xi)$.

Proof. The assertion follows immediately from

$$P(Y \leq t | X < \lambda Y + \xi) = P(Y \leq t,\, X < \lambda Y + \xi)\{P(X - \lambda Y < \xi)\}^{-1}$$

$$= \int_{-\infty}^{t} \Phi(\lambda y + \xi)\, \varphi(y)\, dy\, [\Phi\{\xi(1+\lambda^2)^{-1/2}\}]^{-1}$$

and (2) by differentiation.

Theorem 6
Let T, V be independent random variables, $V \sim \mathscr{N}(0, 1)$, $T \sim \mathscr{T}\{-\xi(1+\lambda^2)^{-1/2}\}$. Then

$$\frac{\lambda}{\sqrt{1+\lambda^2}}\, T + \frac{1}{\sqrt{1+\lambda^2}}\, V \sim \mathscr{SNE}(\lambda, \xi).$$

Proof. A "formal" proof would be the same as the first proof of theorem 1. For a probabilistic proof use lemma 5 and the transformation (5). Since (7) is equivalent to the condition $R < \xi(1+\lambda^2)^{-1/2}$, and since the distribution of $-R$ conditionally on $R < \xi(1+\lambda^2)^{-1/2}$ is that of T, theorem 6 follows from (6).

Note that theorem 1 is a special case of theorem 6.

Corollary 7
Let $Z_{\lambda,\xi} \sim \mathscr{SNE}(\lambda, \xi)$, and let T, V as in the statement of theorem 6. Then

$$E[Z_{\lambda,\xi}^r] = \sum_{j=0}^{r} \binom{j}{r} a^j b^{r-j} E[T^j] E[V^{r-j}]$$

with a, b given in (3).

The problem of computing the moments of the distribution $\mathscr{SNE}(\lambda, \xi)$ thus reduces to the problem of calculating the integral

$$\int_c^\infty t^j \varphi(t)\,dt \qquad (c=-\xi(1+\lambda^2)^{-1/2})$$

which may be done recursively using the equation

$$\frac{d}{dt}\varphi(t)=-t\varphi(t).$$

Finally, theorem 6 implies that the class $\mathscr{SNE}$ (just like $\mathscr{SN}$), after rescaling, is closed with respect to convolution with a centred Normal distribution (property I′ of Azzalini's paper).

References

Azzalini, A. (1985). A class of distributions which includes the Normal ones. *Scand. J. Statist.* **12,** 171–178.

O'Hagan, A. & Leonhard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **63,** 201–202.

Received October 1985, in final form December 1985

Norbert Henze, Universität Hannover, Institut für Mathematische Stochastik, Welfengarten 1, D-3000 Hannover 1, W. Germany