World Applied Sciences Journal 12 (7): 1034-1039, 2011
ISSN 1818-4952

Another Generalization of the Skew Normal Distribution

[1]M.R. Kazemi, [1]H. Haghbin and [2]J. Behboodian

[1]Department of Mathematics, Kazerun Branch, Islamic Azad University, Kazerun, Iran
[2]Department of Mathematics, Shiraz Islamic Azad University, Shiraz, Iran

Abstract: In this paper, we first introduce a new family of distributions, called Minimax Normal distributions, by using Kumaraswamy distribution. Then we use this family to find a generalization of the Balakrishnan skew-normal distribution by the name of skew minimax Normal distribution and we study some of its properties.

Key words: Kumaraswamy distribution . Balakrishnan Skew-Normal distribution . Farlie-Gumble-Morgentern formula . Moment generating function . Conditional probability . Limiting distribution

INTRODUCTION

A random variable X is said to be skew-normal with parameter $\lambda \in R$, written as X~SN (λ), if its density function is

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x), x \in \mathbb{R}$$

where $\phi(x)$ and $\Phi(x)$ denote the N(0,1) density and distribution, respectively, which is introduced by Azzalini [1]. This distribution extends the widely family of Normal distribution by introducing a skewness factor. The skew normal distribution has some applications in different statistical field, like discriminant analysis, regression and graphical models which were discussed by Azzalini and Capitanio [2]. The applications in time series and spatial analysis are given by Genton *et al*. [3] and earlier fragmental works of Aigner *et al*. [4] and Roberts [5] in econometrics and medical studies. Azzalini and Della Valle [6], Azzalini and Capitaneo [2] and Arnold and Beaver [7] discussed the multivariate extension of the distribution; Arellano-Valle *et al*. [8], Gupta and Gupta [9], Jamalizadeh *et al*. [10], Sharafi and Behboodian [11] and Yadegari *et al*. [12] proposed some general classes of the skew-normal distribution.

The paper by Kumaraswamy [13] proposed a new probability distribution for double bounded random processes with hydrological applications. Kumaraswamy's distribution has the following probability density function

$$f(x)=mnx^{m-1}\left(1-x^{m}\right)^{n-1}, 0<x<1 \qquad (1)$$

Recently a natural way of generating families of distributions on some support from a simple starting distribution G with density g is to apply a family of distributions on (0,1), as Jones [14, 15] used this idea to the beta distribution. This family is motivated by the following general class:

$$f(x)=\frac{g(x)G^{a-1}(x)\left(1-G(x)\right)^{b-1}}{B(a,b)} \qquad (2)$$

Eugene *et al*. [16] introduced what is known as the beta normal by taking G in (2) to be cdf of the normal distribution. Now if in (2), beta distribution is replaced by Kumaraswamy's distribution, the resulting density is

$$f(x)=mnG^{m-1}(x)\left\{1-G^{m}(x)\right\}^{n-1}g(x) \qquad (3)$$

Now like beta base distribution family (2) if we take G(.) in (3) to be cdf of another distribution, we have a new family of distributions.

In this paper, by appliying the above disscution, we generalize another aspect of Skew Normal distribution along the line of the Balakrishnan skew normal or Generalized skew normal [17] by reviewing more properties as follows.

In Section 2, we take G in (3) as standard normal distribution. We name this distribution Minimax Normal (MMN) and we find some properties of MMN distribution. In Section 3, a new generalization of Balakrishnan skew normal is introduced by using MMN distribution and it is named Skew Minimax Normal (SMMN); some properties of this distribution is

Corresponding Author: M.R. Kazemi, Department of Mathematics, Kazerun Branch, Islamic Azad University, Kazerun, Iran.

also studied. In Section 4, a multivariate version of Skew Minimax Normal distribution is given.

MINIMAX NORMAL DISTRIBUTION

In this section we define Minimax Normal distribution and study its properties.

Definition 1: A random variable X has MMN (m,n) distribution if its density function is

$$f(x)=mn\Phi^{m-1}(x)\left[1-\Phi^{m}(X)\right]^{n-1}\phi(x),\ x\in R \qquad (4)$$

where m and n are non-negative integer numbers. Another way to construct this distribution involves norma l order statistics.

Consider the following independent and identically distributed random variable from standard normal distribution :

$$\begin{matrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{matrix}$$

then it is easy to show that

$$X=\min_{1\le i\le n}\max_{1\le j\le m}\{X_{ij}\}$$

has MMN(m,n) distribution. Thus, it is reasonable to call this new distribution, Minimax Normal distribution.

Property 1: If X~MMN (m,n), then its cumulative distribution function (cdf) is

$$F_X(x;m,n)=1-\left[1-\Phi^{m}(x)\right]^{n}$$

which can be use to simulate MMN random variable.

Property 2: (Bivariate Generalization): A bivariate generalization of MMN distribution can be derived by using the Farlie-Gumble-Morgentern formula as follows:

$$F_{X_1,X_2}(x_1,x_2)=\left(1-\left[1-\Phi^{m_1}(x_1)\right]^{n_1}\right)\left(1-\left[1-\Phi^{m_2}(x_2)\right]^{n_2}\right)\times \left(1+\alpha\left[1-\Phi^{m_1}(x_1)\right]^{n_1}\left[1-\Phi^{m_2}(x_2)\right]^{n_2}\right)$$

where X_1~MMN(m_1,n_1), X_2~MMN(m_2,n_2) and $-1<\alpha<1$.

Theorem 1: The moment of order k for X~ MMN(m,n) is

$$E\left(X^k\right)=mn\sum_{i=0}^{n-1}\binom{n-1}{i}(-1)^i\times \left[(-1)^k I_{k,(i+1)m-1}+\sum_{j=0}^{(i+1)m-1}\binom{(i+1)m-1}{j}(-1)^j I_{k,j}\right]$$

where

$$I_{k,s}=\frac{1}{2^n\sqrt{2\pi}}\int_0^{\infty} z^k\exp\left(\frac{-z^2}{2}\right)\left[\operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right]^s dz$$

is given by Gupta and Nadarajah [18].

Proof:

$$E\left(X^k\right)=mn\sum_{i=0}^{n-1}\binom{n-1}{i}(-1)^i\int_R x^k\Phi^{(i+1)m-1}(x)\phi(x)dx$$

$$=mn\sum_{i=0}^{n-1}\binom{n-1}{i}(-1)^i \left[(-1)^k I_{k,(i+1)m-1}+\sum_{j=0}^{(i+1)m-1}\binom{(i+1)m-1}{j}(-1)^j I_{k,j}\right]$$

Property 3: If X~MMN (m,n_1) and Y~MMN(m,n_2) are independent random variables, then it can be shown that X|{Y≥X} is distributed as MMN (m,n_1+n_2)

SKEW MINIMAX NORMAL DISTRIBUTION

Balakrishnan [17], as a discussant of Arnold and Beaver [7], generalizes the Azzalini distribution as

$$f_n(x;\lambda)=\frac{\left[\Phi(\lambda x)\right]^n\phi(x)}{C_n(\lambda)},\quad x\in R \qquad (5)$$

where n is a non-negative integer and $C_n(\lambda)$is the normalizing constant. This distribution is known as the Balakrishnan skew normal or Generalized skew normal, denoted by $GSN_n(\lambda)$. A generalization of (5) can be derived as follows:

If X~ N(0,1) and Y~MMN (m,n) are independent random variables, then it can be easily shown that X|{Y≥λX} has a new distribution with the following pdf

$$f_X(x;m,n,\lambda) = \frac{1}{K_{m,n}(\lambda)}\left[1-\Phi^m(\lambda x)\right]^n \phi(x) \qquad (6)$$

where

$$k_{m,n}(\lambda) = P(Y \geq \lambda X) = \int_R \left[1-\Phi^m(\lambda x)\right]^n \phi(x)dx$$

$$= \sum_{i=0}^{n}\binom{n}{i}(-1)^i \int_R \Phi^{im}(\lambda x)\phi(x)dx$$

We will call the pdf (6), Skew Minimax Normal (SMMN) and denote it by $SMMN_{m,n}(\lambda)$

Property 4: It is easy to show that SMMN distribution is a generalized Balakrishnan skew normal distribution (5), because $SMMN_{1,n}(-\lambda) = DSN_n(\lambda)$.

Property 5: If $X:SMMN_{m,n}(-\lambda)$, then $-X:SMMN_{m,n}(-\lambda)$.

Property 6: Let f be the pdf of a $SMMN_{m,n}(\lambda)$, then we have

$$\lim_{\lambda\to-\infty} f_X(x;m,n,\lambda) = 2\phi(x)I(x>0)$$
$$\lim_{\lambda\to+\infty} f_X(x;m,n,\lambda) = 2\phi(x)I(x<0)$$

Theorem 2: The moment of order k for the pdf (6) is as follows,

$$E\left(X^k\right) = \frac{1}{K_{m,n}(\lambda)}\sum_{i=0}^{n}\binom{n}{i}(-1)^i C_{im}(\lambda)E\left(Y_i^k\right)$$

where Y_i: $GSN_{im}(\lambda)$ and $C_j(\lambda)$ is Balakrishnan skew normal constant.

Proof:

$$E\left(X^k\right) = \frac{1}{K_{m,n}(\lambda)}\int_R x^k\left[1-\Phi^m(\lambda x)\right]^n\phi(x)dx$$

$$= \frac{1}{K_{m,n}(\lambda)}\sum_{i=0}^{n}\binom{n}{i}(-1)^i\int_R x^k\Phi^{im}(\lambda x)\phi(x)dx$$

$$= \frac{1}{K_{m,n}(\lambda)}\sum_{i=0}^{n}\binom{n}{i}(-1)^i C_{im}(\lambda)E\left(Y_i^k\right)$$

where

$$C_n(\lambda) = \int_R \Phi^n(\lambda x)\phi(x)dx.$$

is the Balakrishnan skew normal constant.

Property 7: In Theorem 2, if k = 1, then we can present a simple expression for E(X), as follows:

$$E(X) = \frac{\lambda m}{K_{m,n}(\lambda)\sqrt{2\pi\left(1+\lambda^2\right)}}\sum_{i=0}^{n}\binom{n}{i}(-1)^i C_{im-1}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\right)$$

Theorem 3: Let the following sequence of random variables be identically and independently distributed as N(0,1):

$$\begin{matrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{L1} & \cdots & X_{Lm} \end{matrix}$$

And $X\sim SMMN_{m,n}(\lambda)$ be independent from the above sequence. If

$$A = \left\{\min_{1\leq i\leq L}\max_{1\leq j\leq m}\left\{X_{ij}\right\} \geq \lambda X\right\}$$

then X|A is distributed as $SMMN_{m,n+L}(\lambda)$.

Proof:

$$P(X \leq x \mid A) = \frac{\int_{-\infty}^{x} P\left(\min_{1\leq i\leq L}\max_{1\leq j\leq m}\left\{X_{ij}\right\} \geq t\right)f_X(t)dt}{P(A)}$$
$$= \frac{\frac{1}{K_{m,n}(\lambda)}\int_{-\infty}^{x}\left[1-\Phi^m(\lambda t)\right]^{n+L}\phi(t)dt}{P(A)} \qquad (7)$$

It can be shown that

$$P(A) = \frac{K_{m,n+L}(\lambda)}{K_{m,n}(\lambda)}$$

Now, the relation (7) may be re-written as follows:

$$P(X \leq x \mid A) = \frac{\int_{-\infty}^{x}\left[1-\Phi^m(\lambda t)\right]^{n+L}\phi(t)dt}{K_{m,n+L}(\lambda)}$$

and by taking the derivative with respect to x, the proof is complete.

Theorem 4: If $X\sim SMMN_{m,n}(\lambda)$, then its moment generating function is

$$M_X(t)=\frac{1}{k_{m,n}(\lambda)}E\left[\left(1-\Phi^m(\lambda Z+\lambda t)\right)\right]$$

where Z has standard normal distribution.

Proof: Using integration by parts, we would have,

$$\begin{aligned}M_X(t)&=\frac{1}{K_{m,n}(\lambda)}\int_R \exp(tx)\left[1-\Phi^m(\lambda x)\right]^n\phi(x)dx\\&=\frac{1}{K_{m,n}(\lambda)}\int_R\left[1-\Phi^m(\lambda x)\right]^n\phi(x-t)dx\\&=\frac{1}{K_{m,n}(\lambda)}\int_R\left[1-\Phi^m(\lambda y+\lambda t)\right]^n\phi(y)dy\end{aligned}$$

and the proof is complete.

Theorem 5: If X_L~$SMMN_{m,n}(L)$, then $X_L^2\xrightarrow{£}\chi^2_{(1)}$ as $L\to+\infty$, where $\chi^2_{(1)}$ is chi-square random variable with one degree of freedom.

Proof: Let $Y=X_L^2$, then the density of Y is

$$\begin{aligned}f_Y(y;L)&=\frac{1}{\sqrt{y}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y\right)\frac{1}{K_{m,n}(L)}\times\left[\frac{\left[1-\Phi^m\left(L\sqrt{y}\right)\right]^n+\left[1-\Phi^m\left(-L\sqrt{y}\right)\right]^n}{2}\right]\\&=f_{X^2_{(1)}}(y)\left[a_{m,n}(y;L)\right]\end{aligned}$$

with

$$a_{m,n}(y;L)=\frac{1}{K_{m,n}(L)}\left[\frac{\left[1-\Phi^m\left(L\sqrt{y}\right)\right]^n+\left[1-\Phi^m\left(-L\sqrt{y}\right)\right]^n}{2}\right]$$

Since $K_{m,n}(L)\to1/2$ as $L\to+\infty$, then $a_{m,n}(y;L)\to1$ as $L\to+\infty$.

Theorem 6: If X~$SMMN_{m,n}(\lambda)$ and random variable Y which is distributed as chi-squre random variable with degree of freedom r, are independent and

$$W=\frac{X}{\sqrt{\frac{Y}{r}}}$$

then

$$f_W(w)=2f_{T_{(r)}}(w)I(w<0)\quad \text{as } \lambda\to+\infty$$
$$f_W(w)=2f_{T_{(r)}}(w)I(w>0)\quad \text{as } \lambda\to-\infty$$

Proof: By using the joint density of W and Y, then the marginal density of W is

$$f_W(w)=f_{T_{(r)}}(w)\,b_{m,n}(w,\lambda,r)$$

where

$$b_{m,n}(w,\lambda,r)=\frac{\left(1+\frac{w^2}{r}\right)^{\frac{r+1}{2}}}{\Gamma\left(\frac{r+1}{2}\right)\sqrt{2^{r-1}}K_{m,n}(\lambda)}\times\int_0^\infty v^{\frac{r-1}{2}}\exp\left(-\frac{\left(1+\frac{w^2}{r}\right)v}{2}\right)\left[1-\Phi^m\left(\lambda w\sqrt{\frac{v}{r}}\right)\right]^n dv$$

Since

$$\lim_{\lambda\to\infty}K_{m,n}(\lambda)=\frac{1}{2}$$

then we have

$$\lim_{\lambda\to+\infty}b_{m,n}(w,\lambda,r)=2I(w<0)$$

and

$$\lim_{\lambda\to-\infty}b_{m,n}(w,\lambda,r)=2I(w>0)$$

.

Theorem 7: If X~$SMMN_{m_1,n_1}(\lambda)$ and Y~$SMMN_{m_2,n_2}(\lambda)$ are independent and U = X/Y, then

$$f_U(u)=2h(u)I(u>0)\quad\text{as }\lambda\to\pm\infty$$

where h(·) is the standard Cauchy density.

Proof: By using the joint density of U and Y, then the marginal density of U is

$$f_U(u)=h(u)d(m_1,n_1,m_2,n_2,u,\lambda)$$

where

$$d(m_1,n_1,m_2,n_2,u,\lambda)=\frac{\left(u^2+1\right)}{2K_{m_1,n_1}(\lambda)K_{m_2,n_2}(\lambda)}\times\int_{-\infty}^{\infty}|v|\exp\left(-\frac{\left(u^2+1\right)v^2}{2}\right)\left[1-\Phi^{m_1}(\lambda uv)\right]^{n_1}\left[1-\Phi^{m_2}(\lambda v)\right]^{n_2}dv$$

Since

$$\lim_{\lambda\to\infty} K_{m,n}(\lambda) = \frac{1}{2}$$

then we have

$$\lim_{\lambda\to\pm\infty} d(m_1, n_1, m_2, n_2, u, \lambda) = 2\, I(u > 0)$$

MULTIVARIATE SMMN DISTRIBUTION

This Section presents a natural extension of $SMMN_{m,n}(\lambda)$ from the univariate to multivariate case.

Definition 2: (p-dimensional SMMN): Suppose that a random vector X has the density function

$$f_{m,n}(\mathbf{x},\lambda) = \frac{1}{K_{m,n,p}(\lambda)}\left[1-\Phi^m(\lambda'\mathbf{x})\right]^n \varphi_p(\mathbf{x}) \;, \mathbf{x}\in R^p \quad (8)$$

Where

$$K_{m,n,p}(\lambda) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[1-\Phi^m(\lambda'\mathbf{x})\right]^n \varphi_p(\mathbf{x})dx_1\ldots dx_p$$

with $\varphi_p(x)$ as the standard multivariate normal density. Then we say that X has a multivariate SMMN distribution with a vector of skewness parameters $\lambda\in R^p$. For m=1 and n=1 distribution reduces to a multivariate skew normal distribution with skewness parameters $-\lambda$.

Theorem 8: Let $X_1,\ldots,X_p$ be i.i.d. random variables from N(0,1) and Y~MMN (m,n) be independent from this sequence. Now consider the event $A = \{\lambda' X < Y\}$ whith $\lambda\in R^p$ and $X = (X_1,\ldots,X_p)$. Then, X|A has density (8).

Proof: It can be easily shown that the conditional density of (X,Y) given A is

$$f_{\mathbf{X},Y|A}(\mathbf{x},y) = \frac{1}{P(A)}\left[\prod_{i=1}^{p}\varphi(\mathbf{x}_i)\right] f_Y(y) I(\lambda'\mathbf{x} < y)$$

Integrating with respect to y, we have

$$f_{\mathbf{X}|A}(\mathbf{x},y) = \frac{1}{P(A)}\left[\prod_{i=1}^{p}\varphi(\mathbf{x}_i)\right] P(Y > \lambda'\mathbf{x})$$

$$= \frac{1}{P(A)}\left[\prod_{i=1}^{p}\varphi(\mathbf{x}_i)\right]\left[1-\Phi^m(\lambda'\mathbf{x})\right]^n$$

Note that P(A) can be rewritten as follows;

$$P(A) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} P(Y > \lambda'\mathbf{x})\left[\prod_{i=1}^{p}\varphi(\mathbf{x}_i)\right]dx_1\ldots dx_p$$
$$= K_{m,n,p}(\lambda)$$

and the proof is complete.

CONCLUSIONS

This paper introduced a new class of generalized skew-normal distributions, which also includes the Balakrishnan skew-normal distribution. For this distribution, we thoroughly discussed some of its properties during the course of our research. Our procedure is to construct a generalzation of skew-normal distribution based on Kumaraswamy distribution, which has distribution support on interval [0,1].

In stead of kumaraswamy distribution, we will use mixture beta distribution which also has distribution support [0,1], to construct and characterize another new distribution in our future works.

ACKNOWLEGMENT

We would like to thank the editor and reviewers for their constructive comments.

REFERENCES

1. Azzalini, A., 1985. A class of distributions which include the normal. Scandinavian Journal of Statistics, 12: 171-178.
2. Azzalini, A. and A. Capitanio, 1999. Statistical application of the multivariate skew-normal distribution. J. Roy. Statist. Soc., Series B, 61 (3): 579-602.
3. Genton, M.G., L. He and X. Liu, 2001. Moments of skew-normal random vectors and their quadratic forms. Statistics and Probability Letters, 51: 319-325.
4. Aigner, D.J., C.H.K. Lovell and P. Schmidt, 1977. Formulation and estimation of stochastic frontier production function model. Journal of Econometrics, 12: 21-37.
5. Roberts, C., 1966. A correlation model useful in the study of twins. Journal of the American Statistical Association, 61: 1184-1190.
6. Azzalini, A. and A. Dalla Valle, 1996. The multivariate skew normal distribution. Biometrika, 83: 715-726.
7. Arnold, B.C. and R.J. Beaver, 2002. Skew multivariate models related to hidden truncation and/or selective reporting (with discussions) Test, 11: 7-54.

8. Arellano-Valle, R.B., H.W. Gomez and F.A. Quintana, 2004. A new class of skew-normal distribution. Commun. Stat. Theory Methods, 33 (7): 1465-1480.
9. Gupta, R.C. and R.D. Gupta, 2004. Generalized skew-normal model. Test, 13 (2): 501-524.
10. Jamalizadeh, A., J. Behboodian and N. Balakrishnan, 2008. A two-parameter generalized skew-normal distribution. Stat. Probab. Lett., 78: 1722-1726.
11. Sharafi, M. and J. Behboodian, 2008. The Balakrishnan skew-normal density. Stat. Papers, 49: 769-778.
12. Yadegari, I., A. Gerami and M.J. Khaledi, 2008. A generalization of the Balakrishnan skew-normal distribution. Stat. Probab. Lett., 78: 1165-1167.
13. Kumaraswamy, P., 1980. A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46: 79-88.
14. Jones, M.C., 2004. Families of distributions arising from distributions of order statistics (with discussion). Test, 13: 1-43.
15. Jones, M.C., 2009. Kumaraswamys distribution: A beta-type distribution with some tractability advantages. Statistical Methodology, 6: 70-81.
16. Eugene, N., C. Lee and F. Famoye, 2002. Beta-normal distribution and its applications. Commun. Stat. Theory Meth., 31: 497-512.
17. Balakrishnan, N., 2002. Discussion on Skew multivariate models related to hidden truncation and/or selective reporting by B.C. Arnold and R.J. Beaver. Test, 11: 37-39.
18. Gupta, A.K. and S. Nadarajah, 2004. On the moments of the beta-normal distribution. Commun. Stat. Theory Methods, 33: 1-13.