Proc. of the 6th Intl. Symposium on Reliability Engineering and Risk Management (6ISRERM)
31 May – 1 June 2018, Singapore
Editor(s) Xudong Qian, Sze Dai Pang, Ghim Ping Raymond Ong, Kok-Kwang Phoon

VARIOUS TYPES OF THE CUBIC NORMAL DISTRIBUTION AND THEIR APPLICATION IN STRUCTURAL RELIABILITY

Xuan-Yi Zhang[1], Yan-Gang Zhao[2, 3], and Zhao-Hui Lu[1, 3,*]

[1] *School of Civil Engineering, Central South University, 22 Shaoshannan Road, Changsha 410075, China. E-mail:* zhangxuanyi@csu.edu.cn;
[2]*Department of Architecture, Kanagawa University, 3-27-1 Rokkakubashi, Kanagawa-ku, Yokohama 221-8686, Japan. E-mail*: Zhao@kanagawa-u.ac.jp;
[3]*Key Laboratory of Urban Security and Disaster Engineering of Ministry of Education, Beiji ng University of Technology, Beijing 100124, China*
([*]*Corresponding author, Prof. Zhao-Hui Lu, E-mail*: luzhaohui@csu.edu.cn)

Probability distributions of random variables are necessary in engineering practices. However, the distributions of the random variables are usually unknown. It is in this regard that efficient approximation of the distributions is important to accurately quantify practical engineering problems. Generally, probability distributions are determined based on the mean and standard deviation of statistical data, which are not flexible enough to represent the skewness and kurtosis of statistical data. In this paper, we investigate various distribution types of the cubic normal distribution based on the first four moments of the random variables. The explicit expressions of the probability density function and cumulative distribution function of each distribution type are formulated, and the boundaries among different distribution types are presented. Practical examples are applied to demonstrate the efficiency and accuracy of the proposed distribution types in fitting probability distributions of statistical data and structural reliability assessment. We anticipate our assay to be a starting point for precise simulation of the distributions of random variables and thus fulfill more precise evaluation of practical problems.

Keywords: Probability distribution, cubic transformation, the first four moments, statistical data analysis, structural reliability

1 Introduction

Probability distributions of random variables are necessary for engineering practices. Two-parameter (2P) distributions based on the mean and standard deviation of statistical data are often applied to fit these distributions, which may not be appropriate when the skewness and kurtosis of the data are important. Since the bulk of information is often embedded in the first four moments, the distribution of a random variable should be fitted from all first four moments.

Approximating the distribution of a random variable using its moments of finite order is a well-known problem in statistics. The distribution families, such as the Pearson system (1895), the Johnson system and the Burr system (2003) can be used to estimate the distributions of the random variables. Although these systems are flexible, they are practically difficult to implement, especially at the interfaces between different distribution types. Some four-parameter (4P) distributions have been proposed to incorporate the four moments of statistical data, which include the generalized lambda distribution (Ramberg and Schmeiser 1974; Zhao et al. 2006), and S-distribution (Voit 1992). However, their flexibility is attained at a certain price, and the definitions for these distributions are not in a conventional form. Low (2013) proposed a shifted generalized lognormal distribution to overcome these limitations, which needs iterative computation for determining the distribution parameters.

Another method to address this problem is based on the normal transformation, such as the Fisher-Cornish (Fisher and Cornish 1960), Gram-Charlier and Edgeworth series (Wallace 1958). However, these methods specifically generate relatively large error. As alternatives, the Hermite moment model (Winterstein 1988) and cubic normal transformation (Fleishman 1978) are widely applied for data fitting, normal transformation and reliability index in structural reliability. However, some important characteristics of these distributions are still left in the dark, such as the changes in the distribution type with varying combinations of skewness and kurtosis and the boundaries among distinguished types. Without this knowledge, this method cannot be directly applied in practical reliability engineering.

Since all the above-mentioned normal transformation models are cubic polynomials with different definitions of the coefficients, the distribution based on the cubic normal transformation is focused on in this study. The various distribution types in the cubic normal distribution are investigated based on the complete expressions of the cubic normal transformation and its inverse. The boundaries among different types are then graphically presented. The proposed distribution is applied to fit statistical data based on their moments and is found to provide better fitting than existing two- and three-parameter distributions.

2 Various distribution types in the cubic normal distribution

2.1 *Definition of the CDF/PDF of different distribution types*

If the first four moments of a random variable x, i.e., mean (μ_x), standard deviation (σ_x), skewness (α_{3x}), and kurtosis (α_{4x}), are known, the cubic normal transformation can be used to express the standardized random variable $x_s = (x - \mu_x)/\sigma_x$ as follows (Fleishman 1978):

$$x_s = S_u(u) = a_4u^3 + a_3u^2 + a_2u + a_1 \tag{1}$$

where $S_u(u)$ is a third-order polynomial of u; and a_1, a_2, a_3, a_4 are polynomial coefficients determined by setting the first four moments of the left side of Eq. (1) equal to those of the right side (see Appendix A). Based on Eq. (1), the CDF and PDF of x are expressed as:

$$F(x) = \Phi(u) \tag{2a}$$

$$f(x) = \frac{\phi(u)}{\sigma_x(3a_4u^2 + 2a_3u + a_2)} \tag{2b}$$

in which $F(\cdot)$ and $f(\cdot)$ are the CDF and PDF of x, respectively; $\Phi(\cdot)$ and $\phi(\cdot)$ are the CDF and PDF of u, respectively. The expression of u, which is the solution to Eq. **Error! Reference source not found.**) and varies considerably with the order and shape of $S_u(u)$, are listed in Table 1 (Zhao et al. 2018), where the coefficients are formulated as

$$p = \frac{3b - a^2}{9}, q = -\frac{a^3}{27} + \frac{ab}{6} + \frac{a}{2} + \frac{x_s}{2a_4}, \ \Delta = p^3 + q^2, \ \theta = \arccos\left(q/\sqrt[3]{-p}\right) \tag{3}$$

$$J_1^* = \sigma_x a_4(-2|p|^{3/2} - 2q + x_s/a_4) + \mu_x, \ J_2^* = \sigma_x a_4(2|p|^{3/2} - 2q + x_s/a_4) + \mu_x \tag{4}$$

$$J_0 = -(a_2^2/4a_3 + a_3)\sigma_x + \mu_x, \ a = a_3/a_4, \ b = a_2/a_4 \tag{5}$$

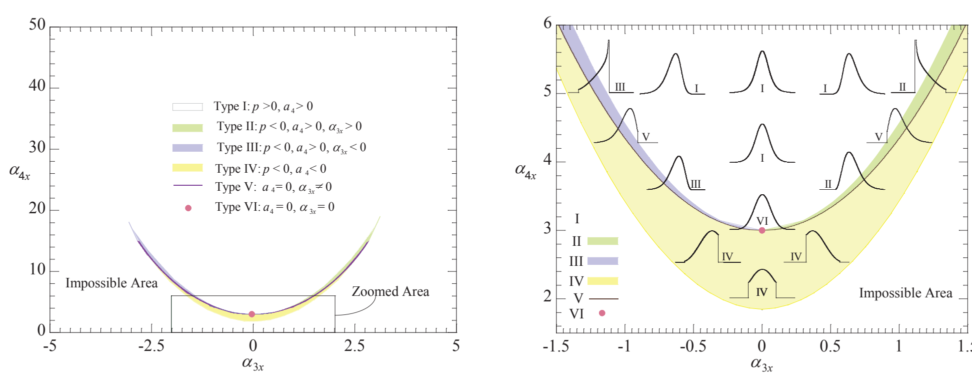
Table 1. Expressions of u

Parameters			u	Range of x	Type
$p \ge 0$			$-p(q+2\sqrt{\Delta})^{-1/3}+(q+2\sqrt{\Delta})^{1/3}-a/3$	$(-\infty, +\infty)$	I
$p < 0$	$a_4 > 0$	$\alpha_{3x} \ge 0$	$2\|p\|^{1/3}\cos(\theta/3)-a/3$	$J_1^* < x < J_2^*$	II
			$-p(q+2\sqrt{\Delta})^{-1/3}+(q+2\sqrt{\Delta})^{1/3}-a/3$	$x \ge J_2^*$	
		$\alpha_{3x} < 0$	$-p(q+2\sqrt{\Delta})^{-1/3}+(q+2\sqrt{\Delta})^{1/3}-a/3$	$x \le J_1^*$	III
			$-2\|p\|^{1/3}\cos[(\theta-\pi)/3]-a/3$	$J_1^* < x < J_2^*$	
	$a_4 < 0$		$-2\|p\|^{1/3}\cos[(\theta+\pi)/3]-a/3$	$J_2^* \le x \le J_1^*$	IV
Not exist	$a_4 = 0$	$\alpha_{3x} > 0$	$[1/4+(a_3/a_2)^2+a_3x_s/a_2]^{1/2}-1/2$	$x \ge J_0$	V
		$\alpha_{3x} < 0$	$[1/4+(a_3/a_2)^2+a_3x_s/a_2]^{1/2}-1/2$	$x \le J_0$	
		$\alpha_{3x} = 0$	x_s	$(-\infty, +\infty)$	VI

According to Table 1, there are six types in the cubic normal distribution, which includes unbounded distributions (Types I and VI), unilaterally bounded distributions (Type II, III and V), and bilaterally bounded distribution (Type IV).

2.2 *Boundaries among different distribution types*

In order to directly obtain the suitable type of the cubic normal distribution using α_{3x} and α_{4x}, the boundaries of each type in the cubic normal distribution are plotted in the α_{3x}-α_{4x} plane, as shown in Fig. 1 (a). The area where $1.5 \le \alpha_{3x} \le 1.5$ and $1.0 \le \alpha_{4x} \le 6$ in Fig. 1(a) is zoomed up. The zoomed area along with some representative PDFs of different types in the cubic normal distribution are plotted in Fig. 1 (b). Figs. 1 (a) and (b) demonstrate that:



(a) Whole range (b) Zoomed area

Fig. 1. Boundaries of the cubic normal distribution.

(a) Type I in the cubic normal distribution covers the largest part in the entire applicable range of the cubic normal distribution.
(b) The applicable areas of Types II and III are symmetric. The applicable area of Type IV is the lowest part in the entire applicable range of the cubic normal distribution. Although the partitions taken by Types II-IV are much smaller than that of Type I, these types cannot be neglected when considering practical engineering problems.
(c) Types V and VI are reduced forms of the cubic normal distribution for $a_4 = 0$, with applicable ranges represented by a bolded curve ($\alpha_{3x} \neq 0$) and a point ($\alpha_{3x} = 0$), respectively.

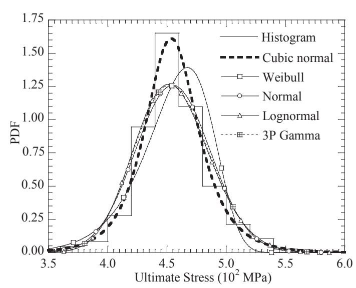
(d) For a fixed value of α_{3x}, with an increase in α_{4x}, the shape of PDF changes from bilaterally bounded form to uniformly bounded form and finally unbounded form. The changing pattern is similar to that of the Pearson system, which indicates that the cubic normal distribution is theoretically reliable. For a fixed value of α_{4x}, the shape of PDF changes from the mn-type to J-type with an increase in $|\alpha_{3x}|$, while the PDF has longer left and right tails for negative and positive α_{3x}, respectively. The cubic normal distribution reflects the characteristics of the skewness and kurtosis well.

3 Application in fitting statistical data

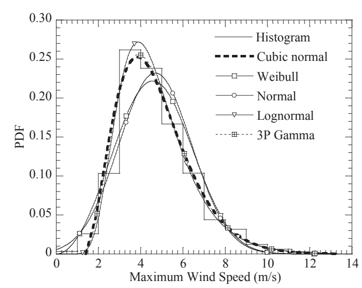
The flexibility of the cubic normal distribution in fitting statistical data is investigated in this section with four sets of data: Sets A is experimental data on the ultimate stress of H-shaped steel (Ono et al 1986); Set B is daily extreme wind speed in Portland at a special station named as ME from 1965 to 1979 (see www.nist.gov/wind); Set C is experimental data on the wind pressure coefficient of a wall-mounted finite-length square cylinder (Wang et al 2015); Set D is measured values of chloride diffusion coefficients gathered in natural structures (Zhang et al 2015). The histograms of the statistical data are presented in Fig. 2, along with the fitted PDFs by using the cubic normal distribution and some commonly used distributions, i.e., Weibull, normal, lognormal, 3P gamma distributions. The sample size, sample moments, and the cubic normal distribution parameters are reported in Table 2. Fig. 2 and Table 2 reveal the following:

Table 2. First four moments of the data and the parameters of the cubic normal distribution

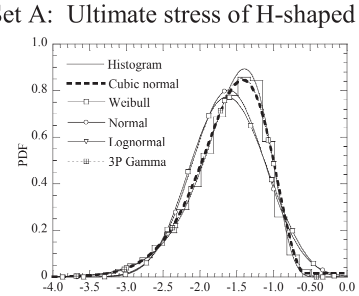
Set	Size	First four moments				Parameters of the cubic normal distribution							
		μ_x	σ_x	α_{3x}	α_{4x}	a_1	a_2	a_3	a_4 ($\times 10^{-3}$)	p	J_1^*	J_2^*	Type
A	1932	4.549	0.317	0.153	6.037	-0.017	0.782	0.017	68.05	11.46	---	---	I
B	5478	4.687	1.719	0.922	4.153	-0.156	0.974	0.156	0.462	-35720.9	8927523	-6289.22	II
C	19000	-1.609	0.497	-0.750	3.755	0.126	0.984	-0.126	0.04	-2705734	23012.22	-1.7×10^9	III
D	50	4.328	2.114	0.467	2.854	-0.90	1.058	0.90	-22.66	-51.95	177.095	-432.236	IV



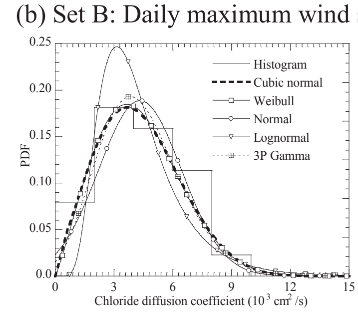
(a) Set A: Ultimate stress of H-shaped steel



(b) Set B: Daily maximum wind speed



(c) Set C: Experimental wind pressure coefficient



(d) Set D: Chloride diffusion coefficients

Fig. 2. Comparison among popular distributions and the cubic normal distribution in fitting actual data.

(a) Among the five distributions, the PDFs of the 2P distributions, i.e., Weibull, normal and inverse normal distributions, have the greatest differences from the histogram of the statistical data, except for set D, where the Weibull distribution performs as well as the cubic normal distribution. In particular, when the value of the statistical data is negative, e.g., set C, Weibull and lognormal distributions cannot be directly applied due to their limitations, and the corresponding PDFs are obtained by adopting the absolute value of the statistical data.
(b) The 3P gamma distribution fits the histogram of the statistical data much better than the 2P distributions when the kurtosis is around 3, as shown in Fig. 2(c). However, when the kurtosis is large, the PDF of the 3P gamma distribution is considerably different from those of the histograms in Fig. 2(a).
(c) The cubic normal distribution fits the histogram much better than the 2P distributions and the 3P gamma distribution, and the results of the cubic normal distribution are in close agreement with the histograms of the statistical data for all four cases considered in this study.

4 Application in structural reliability assesment

In engineering practices, the cubic normal distribution can be used to simulate the CDFs/PDFs of random variables with the aid of Eqs. (2a) and (2b) and Table 1. Reliability analysis can then be further conducted using general analysis methods such as the first-order reliability method (FORM) and second-order reliability method (SORM).

The second example considers a composite beam with 19 independent variables, as shown in Fig. 3. To keep the composite beam in the safe domain, the allowable strength should be larger than the maximum stress, and thus the performance function is determined as (Xiao et al. 2014)

$$G(\mathbf{X}) = S_a - \frac{\left[\dfrac{\sum_{i=1}^{6} P_i\left(L-L_i\right)}{L}L_3 - P_1\left(L_2-L_1\right)-P_2\left(L_3-L_1\right)\right]\left[\dfrac{0.5A_0B_0^2+\dfrac{E_a}{E_w}D_0C_0\left(B_0+D_0\right)}{A_0B_0+\dfrac{E_a}{E_w}D_0C_0}\right]}{\dfrac{1}{12}A_0B_0^3+A_0B_0\left[\dfrac{0.5A_0B_0^2+\dfrac{E_a}{E_w}D_0C_0\left(B_0+D_0\right)}{A_0B_0+\dfrac{E_a}{E_w}D_0C_0}-0.5B_0\right]^2+\dfrac{1}{12}\dfrac{E_a}{E_w}C_0D_0^3+\dfrac{E_a}{E_w}C_0D_0\left[0.5D_0+B_0-\dfrac{0.5A_0B_0^2+\dfrac{E_a}{E_w}D_0C_0\left(B_0+D_0\right)}{A_0B_0+\dfrac{E_a}{E_w}D_0C_0}\right]^2} \tag{6}$$

where S_a is the allowable strength; L_1, L_2,…, L_6 and L are the measured corresponding length from the left node; E_a and E_w are the Young's moduli; P_1, P_2,…, P_6 are the applied loads at six different locations along the beam; A_0, C_0, and B_0, D_0 are the width and height of the beam, respectively; A_0 is a design variable with the lower and upper bounds of 100 mm and 120 mm, respectively. Detailed information of the random variables is given in Table 3.

Because all the random variables in Eq. (6) have known CDFs/PDFs, the reliability index can be readily obtained using FORM/SORM/MCS. To investigate the efficiency of the proposed reliability method, the CDFs/PDFs of all the random variables are assumed to be unknown, and only their first four moments are assumed to be known. With the first four moments, the parameters of the cubic normal distribution can be readily obtained, as also listed in Table 3. Then, the CDFs/PDFs of the random variables can be easily approximated using Eqs. (2a) and (2b) and Table 1. In this example, the distribution of S_a is Type II in the cubic normal distribution, the distribution of L_i (i = 1 ~ 6) and L are Type VI in the cubic normal distribution, and those of the other random variables are Type I in the cubic normal distribution.

The first-order reliability indices, β_{FORM}, obtained using the assumed distribution and the cubic normal distribution as well as the reliability indices obtained by MCS with 1×10^7 samples (the COVs of the MCS results vary from 0.315% to 0.502%) are depicted in Fig. 4. The first-order reliability indices obtained from the cubic normal distribution coincide with those obtained by the assumed distribution. However, the first-order reliability indices differ greatly from those obtained by MCS due to the strong nonlinearity of the performance function in Eq. (6).

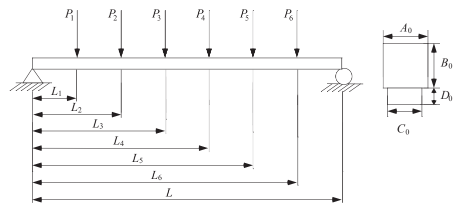


Fig. 3. A composite beam in Example 2

Table 3. The first four moments and the parameters of the cubic normal distribution of random variables in Example 2

Variables	Distribution	The first four moments				The parameters of the cubic normal distribution						
		μ_x	σ_x	α_{3x}	α_{4x}	a_2	a_3	a_4	p	J_1^*	J_2^*	Type
S_a	Gamma	25	2.5	0.200	3.06	0.9981	0.0333	0.0003	-1210.9	2.7726	25.251	II
B_0 (mm)	Lognormal	200	0.2	0.003	3	1	0.0005	1.67×10^{-7}	3×10^6	---	---	I
C_0 (mm)	Lognormal	80	0.2	0.0075	3	1	0.0013	1.04×10^{-6}	479998	---	---	I
D_0 (mm)	Lognormal	20	0.2	0.030	3	1	0.0050	1.67×10^{-5}	29998.5	---	---	I
L_1 (mm)	Normal	200	1	0	3	1	0	0	---	---	---	VI
L_2 (mm)	Normal	400	1	0	3	1	0	0	---	---	---	VI
L_3 (mm)	Normal	600	1	0	3	1	0	0	---	---	---	VI
L_4 (mm)	Normal	800	1	0	3	1	0	0	---	---	---	VI
L_5 (mm)	Normal	1000	1	0	3	1	0	0	---	---	---	VI
L_6 (mm)	Normal	1200	1	0	3	1	0	0	---	---	---	VI
L (mm)	Normal	1400	1	0	3	1	0	0	---	---	---	VI
P_i (kN)	Gumbel	15	1.5	1.140	5.4	0.897	0.168	0.0242	20.954	---	---	I
E_a (GPa)	Lognormal	70	7	0.301	3.162	0.9925	0.0497	0.0017	298.485	---	---	I
E_w (GPa)	Lognormal	8.75	0.875	0.301	3.162	0.9925	0.0497	0.0017	298.485	---	---	I

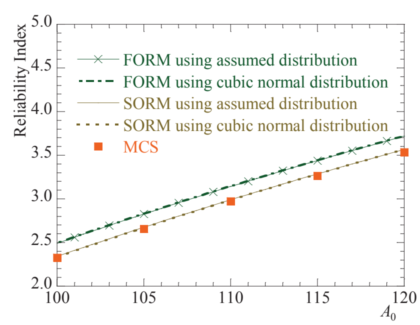


Fig. 4. Reliability indices obtained using the assumed distribution and the cubic normal distribution in Example 2

Using the point-fitting SORM (Zhao and Ono 1999), the second-order reliability indices, β_{SORM}, obtained using both the assumed distribution and the cubic normal distribution are also depicted in Fig. 4. The second-order reliability indices obtained using the cubic normal distribution are nearly identical to those obtained using the assumed distribution and are in close agreement with the MCS results.

5 Conclusion

This study investigates various distribution types of the cubic normal distribution, which contains unbounded, left-bounded, right-bounded and bilaterally bounded distributions. Explicit expressions of different distribution types have been proposed. The boundaries among distinguished types and the applicable range of the random variable for each type are defined. Numerical examples demonstrate that: (1) The cubic normal distribution has rich flexibility in fitting probability functions of statistical data such as ultimate stress of H-shaped steel, average wind speeds, wind pressure coefficient, and chloride diffusion coefficients; (2) Structural reliability analysis using first- and second-order reliability method is enough accurate when the probability distributions of the basic random variables are approximated by the cubic normal distributions.

Acknowledgments

The study is partially supported by the National Natural Science Foundation of China (Grant No.: 51738001, U1434204, and 51421005), the scholarship from China Scholarship Council (Grant No.: 201706370095) and the Fundamental Research Funds for the Central Universities of Central South University (Grant no.: 1053320180507). The supports are gratefully acknowledgement.

References

Ang, AH-S., Tang, W. H., *Probability Concepts in Engineering Planning and Design. Vol I: Basic Principles*, New York: John Wiley & Sons, 1975.

Fisher, R. A., Cornish, E. A., The percentile points of distributions having known cumulants, *Technometrics.* 2, 209-225, May, 1960.

Fleishman. A method for simulating non-normal distributions, *Psychometrika*. 43(4), 521-532, Dec, 1978.

Kleiber, C., Kotz, S., *Statistical size distributions in economics and actuarial sciences*, Hoboken, New Jersey: Wiley; 2003.

Low, Y. M., A new distribution for fitting four moments and its applications to reliability analysis, *Struct. Saf*., 42, 12-25, May, 2013.

Ono, T., Idota, H., Kawahara, H., A statistical study on resistances of steel columns and beams using higher order moments, *J. Struct. Constr. Eng.* AIJ, 370, 19-37, May, 1986. (in Japanese)

Pearson, K., Contributions to the mathematical theory of evolution, II: skew variation in homogeneous material, *Philos. Trans. R. Soc. London (A)*., 186, 343-414, Jan, 1895.

Ramberg, J. S., Schmeiser, B.W., An approximate method for generating asymmetric random variables, *Commun. ACM*., 17(2), 78-82, Feb, 1974.

Voit, E. O., The S-distribution a tool for approximation and classification of univariate, unimodal probability distributions, *Biometrical. J*., 34(7), 855-78, Jan, 1992.

Wallace, D. L., Asymptotic approximations to distributions, *Ann. Math. Statist.* 29(3), 635-654, Sep, 1958.

Wang, H. F., Zhao, X. Y., He, X. H., Zhou, Y. Effects of oncoming flow conditions on the aerodynamic forces on a cantilevered square cylinder. *J. Fluid Struct*., 75, 140-157, Nov, 2017.

Winterstein, S. R., Nonlinear vibration models for extremes and fatigue, *J. Eng. Mech*., 114(10), 1772-1790, Oct., 1988.

Xiao, N.C., Li, Y.F., Yang, Y.J., Yu, L., Huang, H.Z. A novel reliability method for structural systems with truncated random variables. *Struct. Saf.*, 50, 57-65, Sept., 2014.

Zhao, Y. G., Lu Z. H., Ono T., 4P-Lambda distribution and its applications to structural reliability assessment, *J. Struct. Constr. Eng.* AIJ, 604, 47-54, Aug, 2006.

Zhang, X. G., Wang, J., Zhao, Y. G., Tang, L. P., Xing, F., Time-dependent probability assessment for chloride induced corrosion of RC structures using the third-moment method, *Constr. Build. Mater.*, 76, 232-244, Feb., 2015.

Zhao, Y. G., Zhang, X. Y., Lu, Z. H., Complete Monotonic Expression of the Fourth-Moment Normal Transformation for Structural Reliability. *Comput. Struct*., 196, 186-199, Dec., 2018.

Zhao, Y.G., Ono, T. New approximations for SORM: Part II. *J. Eng. Mech.*, 125(1): 86-93, Jan.,1999.