Structural Safety 61 (2016) 57–66

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

Translation random field with marginal beta distribution in modeling material properties

Yong Liu *, Ser-Tong Quek, Fook-Hou Lee

Department of Civil and Environmental Engineering, National University of Singapore, Block E1A, #07-03, No. 1 Engineering Drive 2, Singapore 117576, Singapore

ARTICLE INFO

Article history:
Received 19 November 2015
Received in revised form 17 April 2016
Accepted 18 April 2016
Available online 28 April 2016

Keywords:
Beta distribution
Translation random field
Material property
Monte-Carlo simulations
Correlation structure

ABSTRACT

For modeling material properties having a bounded range, the beta distribution may be adopted as the marginal distribution of a second-order non-Gaussian random field. Three aspects related to the simulation of such random field are discussed in this study. First, an unbiased and consistent estimator for the lower (and upper) bound of the beta distribution based on sample data is proposed. This estimator is shown to be generally more efficient than that given by the method of moments. Second, a simple explicit function relating the auto-correlation function of the non-Gaussian random field to that of the underlying Gaussian field is proposed. The relationship facilitates control on the scale of fluctuation of the non-Gaussian field. Third, an algorithm is proposed for generating random fields with an approximate marginal beta distribution and a prescribed cross-correlation, where the latter can range from −1 to 1. Numerical examples are given to illustrate the effectiveness and efficiency of each of the three aspects. The estimation of the lower bound of material property is exemplified through field data from a real project.

1. Introduction

The presence of uncertainties in the material properties has been acknowledged by the engineering community and accounted for in codes of practice [1,2]. By treating the material properties as random variables or in some cases as random fields, statistics-based reliability analysis and the specification of characteristic values are techniques adopted to handle uncertainties in a consistent manner [3–6].

The normal and lognormal distributions are widely used to describe random variables in practice [6] due to their simplicity and well-established properties, notwithstanding that often the actual distributions are hard to establish due to the limited amount of data. However, there are engineering variables that have obvious lower and upper limits and there are instances where large volumes of data are available. For the latter, high order moments (e.g. skewness and kurtosis) of a random variable may even be estimated with fairly good accuracy. Some distributions are only uniquely defined by higher order moments, such as the maximum entropy distribution [7–9], Hermite polynomials of normal variates [10] and Johnson's system of frequency distributions [11]. Recently, Low [12] developed a four-parameter distribution to reflect the first four moments of sampled data.

The current study adopts the beta distribution as it not only can capture the first four moments of a random variable, but is bounded. In addition, it has well-established properties for computer-aided simulations. Many publications [3,4,13,14] have shown the flexibility of the beta distribution in reflecting the variability of material properties. Harrop-Williams [15] analytically showed that the strength parameters (cohesion and the tangent of friction angle) of uniform soils follow the beta distribution. Therefore, it is of practical interest to establish methods to efficiently and accurately estimate the parameters of beta distribution from sampled data.

Unlike the normal and lognormal distributions which are completely defined by the sample mean and variance, the beta distribution requires additional information, such as the distribution bounds or higher order moments. Cooke [16] proposed a method to estimate the bounds of general random variables. Based on Cooke's method, He [17] proposed an iterative algorithm to estimate the bounds of beta-distributed random variables. This algorithm was simplified by Liu et al. [18]. However, both the original and simplified iterative algorithms have certain conditions for convergence, and the iterative process may lead to biased estimators. Another commonly used approach [3,19] is to fix the range of the distribution to six sample standard deviations centered at the

* Corresponding author.
E-mail addresses: ceeliuy@gmail.com (Y. Liu), st_quek@nus.edu.sg (S.-T. Quek), ceeleefh@nus.edu.sg (F.-H. Lee).

http://dx.doi.org/10.1016/j.strusafe.2016.04.001
0167-4730/

sample mean, which may not reflect correctly the skewness of the distribution. In this paper, a non-iterative approach is proposed to estimate the bounds of a beta variable. It will be shown that this leads to unbiased and consistent estimators. The approach will be compared with the *method of moments* [20,21], where the bounds are obtained through the first four moments of sampled data.

In the case of a medium with material property that fits the marginal beta distribution, a second-order beta random field [22] may be considered. One way to generate such beta field is to translate from an underlying Gaussian field [23]. However, the translation changes the correlation structure [24]. Yamazaki and Shinozuka [25] proposed a method to rectify the simulated correlation structure by iteratively updating the power spectral density function (PSDF) (which is the Wiener-Khintchine transform of correlation function) of the underlying Gaussian field until a desired PSDF of the translation field is achieved. This updating method has been extensively developed [26–28]; in particular, Shields et al. [28] proposed an iterative translation approximation method to speed up the calculation. However, the iteration process in simulating random field might be time-consuming [29,30], where additional effort in generating the underlying Gaussian field is needed to match the PSDF. Matching the PSDF may preserve the form of the correlation function; however, the exact scale of fluctuation (SOF) may not be realized. This paper proposes an alternative method where the SOF can match exactly the prescribed value without iteration, assuming that the exact form of the auto-correlation function is not as critical in stochastic finite element analysis. On the other hand, based on Grigoriu's [10] work, the lower bound of correlation of a translation random field is generally greater than -1 [31]. Some material properties indeed have negative correlation (close to -1), such as the friction angle and effective cohesion of soils [13,32]. The lower bound in correlation may be necessary and sufficient to reflect the cross-correlation in such kinds of material properties. To circumvent this difficulty, a method based on Cholesky decomposition [33] is proposed to generate a beta field with a prescribed cross-correlation ranging from -1 to 1.

2. Beta distribution

2.1. Probability density function

A random variable X distributed in the interval $[a, b]$ is said to follow the beta distribution, if its probability density function (PDF), $f_X(x;\ \alpha,\ \beta)$, satisfies:

$$f_X(x;\alpha,\beta) = \begin{cases} \frac{(x-a)^{\alpha-1}(b-x)^{\beta-1}}{\mathbf{B}(\alpha,\beta)(b-a)^{\alpha+\beta-1}}, & a \leqslant x \leqslant b \\ 0, \ \text{otherwise} \end{cases} \tag{1}$$

in which a and b are the lower and upper bounds of X, respectively; α and β are the shape parameters, which are positive real numbers; $\mathbf{B}$ is the beta function [20]. The cumulative distribution function (CDF) of X, $F_X(x;\ \alpha,\ \beta)$, can be obtained from Eq. (1) via integration.

2.2. Estimation of bounds

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a beta distribution $F_X(x;\ \alpha,\ \beta)$. For simplicity, this sequence is assumed to be an ordered statistics with X_1 and X_n being the minimum and maximum of the sequence, respectively, and the sample size n is reasonably large (say larger than 20). The shape parameters (α, β) are assumed to be known, or can be obtained by the method outlined in Oboni and Bourdeau [34] or estimated from the first four moments of the sampled data [21]. It can be shown that the following estimator $\widehat{b}$ gives an unbiased and consistent estimation [35] of the upper bound b:

$$\widehat{b} = X_n + \beta \cdot (X_n - X_{n-1}) \tag{2}$$

Proof:

As the maximum of a sample, the CDF of X_n is given by:

$$F_{X_n}(x;\alpha,\beta) = [F_X(x;\alpha,\beta)]^n \tag{3}$$

For a reasonably large n, Eq. (3) leads to the Type III extreme value distribution and may be written as [16,36]:

$$F_{X_n}(x;\alpha,\beta) = \exp\left\{-\left(\frac{b-x}{b-\kappa_n}\right)^{\beta}\right\} \tag{4}$$

in which

$$\kappa_n = F_X^{-1}(1-1/n;\alpha,\beta) \tag{5}$$

The expectation of X_n can be calculated as:

$$E\{X_n\} = \int_{-\infty}^{b} x\mathrm{d}F_{X_n}(x;\alpha,\beta) \tag{6}$$

The right-hand side of Eq. (6) can be integrated by parts:

$$E\{X_n\} = b - \int_{-\infty}^{b} F_{X_n}(x;\alpha,\beta)dx \tag{7}$$

Substituting Eq. (4) into Eq. (7) and integrating the CDF gives:

$$E\{X_n\} = b - \frac{b-\kappa_n}{\beta}\Gamma\left(\frac{1}{\beta}\right) \tag{8}$$

in which $\Gamma(\cdot)$ is the gamma function [20]. Similarly, the expectation of X_{n-1} can be obtained as:

$$E\{X_{n-1}\} = b - (b-\kappa_n)\frac{\Gamma(1/\beta+2)}{\Gamma(2)} \tag{9}$$

Since $\Gamma(1/\beta+2) = (1/\beta+1)\cdot\Gamma(1/\beta+1)$ and $\Gamma(2) = 1$, the expectation of $\widehat{b}$ in Eq. (2) can therefore be calculated as:

$$E\{\widehat{b}\} = b \tag{10}$$

Thus, $\widehat{b}$ is an unbiased estimator of the upper bound b. For the extreme case where n tends to an infinitely large value, both X_n and X_{n-1} converge in probability towards the upper bound b, which implies that $\widehat{b}$ equals to b under this extreme case. Thus, $\widehat{b}$ is also a consistent estimator [35] of the upper bound b. ■

Likewise, it can be shown that the estimator $\widehat{a}$:

$$\widehat{a} = X_1 - \alpha \cdot (X_2 - X_1) \tag{11}$$

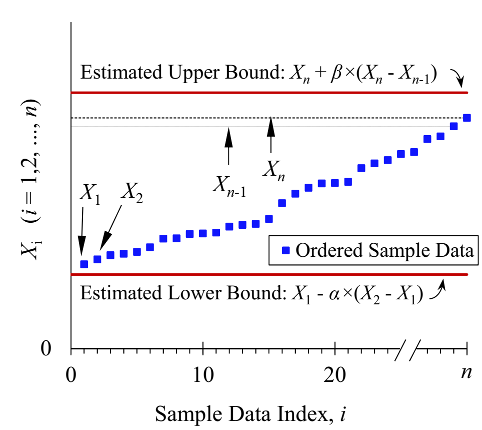

Fig. 1. Estimation of bounds based on ordered sample data.