



Approximation of continuous random variables for the evaluation of the reliability parameter of complex stress–strength models

Alessandro Barbiero¹ · Asmerilda Hitaj²

Accepted: 24 February 2021
© The Author(s) 2021

Abstract

In many management science or economic applications, it is common to represent the key uncertain inputs as continuous random variables. However, when analytic techniques fail to provide a closed-form solution to a problem or when one needs to reduce the computational load, it is often necessary to resort to some problem-specific approximation technique or approximate each given continuous probability distribution by a discrete distribution. Many discretization methods have been proposed so far; in this work, we revise the most popular techniques, highlighting their strengths and weaknesses, and empirically investigate their performance through a comparative study applied to a well-known engineering problem, formulated as a stress–strength model, with the aim of weighting up their feasibility and accuracy in recovering the value of the reliability parameter, also with reference to the number of discrete points. The results overall reward a recently introduced method as the best performer, which derives the discrete approximation as the numerical solution of a constrained non-linear optimization, preserving the first two moments of the original distribution. This method provides more accurate results than an *ad-hoc* first-order approximation technique. However, it is the most computationally demanding as well and the computation time can get even larger than that required by Monte Carlo approximation if the number of discrete points exceeds a certain threshold.

Keywords Approximation · Cumulative distribution function · Discretization · Moment equalization · Monte Carlo simulation

✉ Alessandro Barbiero
alessandro.barbiero@unimi.it
Asmerilda Hitaj
asmerilda.hitaj@uninsubria.it

¹ Department of Economics, Management and Quantitative Methods, Università degli Studi di Milano, Via Conservatorio 7, 20122 Milan, Italy

² Department of Economics, Università degli Studi dell’Insubria, Via Monte Generoso 71, 21100 Varese, Italy

1 Introduction

Product design is a crucial stage of product lifecycle if one aims at ensuring high quality in a product, both from an engineering and marketing perspective. Product quality can be measured in terms of the reliability of the product given a defined environment or a defined mission time. Under an engineering viewpoint, a product, should it be a component or a system consisting of many components or elements that interact with one another, can be regarded as having its own strength, which is random in nature and is usually assumed to be continuous. The product is subject to some stress, which can be modeled as a random variable (rv) as well, and works correctly as long as the strength overcomes the stress: the probability of this random event occurring is called the reliability parameter (of the product, system, component, etc). Since stress and strength are often functions of elementary stochastic factors and the form of these functions is usually very complex, it descends that finding their exact statistical distribution, and then the value of the reliability parameter, is at least cumbersome if not actually impossible through customary integration techniques. It is standard practice to carry out Monte Carlo simulations in order to find this value numerically. A convenient alternative solution to this impasse comprises discretization, i.e., substituting the probability density function (pdf) of each continuous rv involved through a finite number of points with assigned probabilities. Then, an approximate value of reliability can be recovered by enumeration.

Many approximation-by-discretization methods have been proposed in the literature so far, which often differ from one another for their range of application and ultimate scope. We can roughly divide them into two classes. A first class of methods (which we can name moment-matching techniques) are based on the equalization of the first moments of the original continuous variable; among them, a recent proposal that derives the discrete approximating rv as the solution to an optimization problem, which preserves expected value and variance. A second class of method are based on the (preservation of the) survival function at the discrete points of the approximating distribution; within this class, a proposal of discretization, based on the minimization of a distance between the cumulative distribution functions of the original and approximating rv's, is worth being exploring.

In this work, a comparative study, applied to a practical and well-known engineering problem, will empirically investigate the performance of these methods, by involving symmetrical and asymmetrical probability distributions, and varying the number of approximating discrete points; a comparison to an *ad hoc* first-order approximation technique will be also performed and some practical advice on their mindful use will be provided.

The paper is organized as follows. In Sect. 2, we introduce the stress–strength model and the problem of computing the reliability parameter for complex stress–strength models, also resorting to approximation or discretization techniques, which avoid multivariate numerical integration. In Sect. 3, the concept of discretization is framed, several univariate discretization techniques are presented according to the above-mentioned classification, with the aim of sorting out their utility, strengths and weaknesses. Section 4 illustrates a comparative closeness study with an application to an engineering problem. Some final comments are left for Sect. 5.

2 Computing the reliability parameter of complex stress–strength models

In simple terms, a stress–strength model consists of an item (or, more generally, a component or a system) that can be modelled as having an intrinsic strength X , operating under a stress Y . The magnitudes of both stress on, and strength of, an item are usually modelled as continuous rv's and assumed to be independent and follow some completely or partially known distribution. The reliability (or reliability parameter) of the item is defined as the probability that it properly works, i.e., the probability that the strength exceeds the stress: $R = P(X > Y)$.

In complex stress–strength models, stress and strength are two functions of several stochastic factors. In more detail, let X be a known function of n sub-factors of strength: $X = \phi_1(X_1, X_2, \dots, X_n)$, and Y a known function of m sub-factors of stress: $Y = \phi_2(Y_1, Y_2, \dots, Y_m)$; where X_l and Y_j ($l = 1, \dots, n, j = 1, \dots, m$) are mutually independent and follow each a fully specified continuous distribution. Denoting with $f_1(x)$ and $f_2(y)$ the pdf of X and Y , then R is given by

$$R = \int_{-\infty}^{+\infty} dx \int_{-\infty}^x f_1(x) f_2(y) dy, \quad (1)$$

since X and Y turn out to be statistically independent. As said in the introduction, computing the exact value of R given by Eq. (1) in this context through the common analytic techniques is not an easy task, because deriving $f_1(x)$ and $f_2(y)$ is not straightforward, since numerical integration methods (Stroud, 1971; Cools, 2002) become impracticable even for small values for n and m ; the approximation of a continuous rv through discretization represents an alternative solution, along with *ad-hoc* approximation methods and Monte Carlo (MC) simulation for the evaluation of the reliability parameter.

In order to compute R through MC simulation, a large number S of random values has to be produced for each stochastic factor. For each MC run s , from 1 to S , $x_{1s}, \dots, x_{ns}, y_{1s}, \dots, y_{ms}$ have to be independently drawn from the known distributions of $X_1, \dots, X_n, Y_1, \dots, Y_m$; then the values $x_s = \phi_1(x_{1s}, \dots, x_{ns})$ and $y_s = \phi_2(y_{1s}, \dots, y_{ms})$ are computed. R is then approximated by the proportion of x_s values greater than the corresponding y_s :

$$R_{sim} = \sum_{s=1}^S \mathbb{1}[x_s > y_s] / S \quad (2)$$

where $\mathbb{1}[E]$ is the indicator function, having value 1 if E is true, 0 if E is false. Due to the law of large numbers, we know that R_{sim} is a consistent and asymptotically normal estimator of the “true” value of the parameter R . In fact, every indicator $\mathbb{1}[x_s > y_s]$ can be regarded as a realization of an independent and identical copy of a Bernoulli rv with success probability $R = P(X > Y)$, and R_{sim} is then the sample mean of S independent and identically distributed (i.i.d.) rvs with expected value R and variance $R \cdot (1 - R)$, so that the law of large numbers can be applied. By considering a large value for the number of simulations S , we expect the actual value of R_{sim} is (on average) very close to R .

We will now briefly recall first- and second-order reliability methods, FORM and SORM (Der Kiureghian, 2004). FORM has been widely used in structural reliability estimation applications. If we define the random vector \mathbf{X} as the juxtaposition of strength and stress stochastic factors, $(X_1, \dots, X_n, Y_1, \dots, Y_m)$ and the limit state function (LSF) as $g(\mathbf{X}) = X - Y = \phi_1(X_1, \dots, X_n) - \phi_2(Y_1, \dots, Y_m)$, we have that $g(\mathbf{X}) \leq 0$ defines the “failure region” and $g(\mathbf{X}) > 0$ defines the “safe region”. The failure probability, i.e.,

$P(X \leq Y)$, can be written as

$$P_f = \int_{g(\mathbf{X}) \leq 0} f_1(x_1, \dots, x_n) f_2(y_1, \dots, y_m) dx_1 \dots dx_n dy_1 \dots dy_m$$

which, as anticipated, is not easily derivable. FORM involves Taylor expansion of the LSF, i.e., its linearization, which is not performed around the mean value of the function, but at a point that is called the “most probable failure point” (MPP) or “design point”. The process starts with the transformation of the non-normal variables to standard normal variables with zero mean and unit variance, using the Rosenblatt transformation: from the “original space” of (non-normal) variables, one moves to the “reduced space” of corresponding standard normal variables. The target is to find the MPP, defined as the point on the limit state surface ($g(\mathbf{X}) = 0$) that defines the minimum distance from the origin in the space of the reduced variables. The shortest distance between the failure surface and the origin in the space of the reduced variables is called the “reliability index” β . The MPP, and then β , can be found by using an iterative procedure; once β is available, the failure probability of the system can be approximated by $P_f \approx \Phi(-\beta)$, where Φ is the cumulative distribution function (cdf) of the standard normal rv, and then the approximate value of the reliability parameter R is calculated as $1 - P_f$.

FORM is based on the first order approximation of the LSF. Of course, the results are the more accurate the closer the cdf’s are to the normal and the less the LSF deviates from a linear form. Clearly, if all the rvs involved are normal and the LSF is linear, then the provided value is exact: $P_f = \phi(\beta)$. If one more term is included in its Taylor expansion, we have a “second order” approximation (SORM), which includes also the second order derivatives of $g(\mathbf{X})$, and is expected to provide more accurate estimates of P_f , especially when P_f is not close to zero.

An approximate value of R can be recovered by constructing a discrete counterpart for each stochastic factor (discretization). Whatever discretization approach is applied, the reliability parameter R of the system can be then approximately computed as:

$$R \approx \sum_{x_{d,1}} \dots \sum_{y_{d,m}} \left[\prod_{l=1}^n P(X_{d,l} = x_{d,l}) \right] \cdot \left[\prod_{j=1}^m P(Y_{d,j} = y_{d,j}) \right] \cdot \mathbb{1}[\phi_1(x_{d,1}, \dots, x_{d,n}) > \phi_2(y_{d,1}, \dots, y_{d,m})], \quad (3)$$

where the sums extend over all possible $x_{d,l}$ and $y_{d,j}$, which are the values that the discretized versions of $X_l (X_{d,l})$ and $Y_j (Y_{d,j})$, $l = 1, \dots, n$, $j = 1, \dots, m$, can assume. If, for example, X and Y are functions of $n = 2$ and $m = 3$ random factors, respectively, and all the random factors are discretized through $k = 4$ points, the approximation consists of the construction a table of $k^n \cdot k^m = 4^2 \cdot 4^3 = 1024$ combinations by which an approximate value of reliability can be easily computed by enumeration.

This discretizing approach was first explored by Taguchi (1978), who suggested a factorial experiment approach, based on a discretization of the continuous rv modeling stress into a 3-point discrete distribution, with equal probabilities. D’Errico and Zaino (1988) provided a modification to Taguchi’s approach, by suggesting a 3-point discrete distribution that retains the first five raw moments of a normal rv by using unequal probabilities. English et al. (1996) extended this method up to 6 points and applied it in the solution of four engineering problems. Roy and Dasgupta (2001), Roy and Dasgupta (2002), Barbiero (2012) employed their own discretization techniques, which will be described in the next section, to evaluate the reliability parameter of the complex stress–strength models investigated in English et al. (1996).

Apart from stress–strength models, engineering applications provide a myriad of non-linear transformations that are suitable for applying the discretization method (Luceno, 1999); however, the common basic underlying idea is to replace pdf of continuous rvs with discrete approximations, so that the parameter object of the study can be recovered by enumeration rather than inversion theorem, convolution integrals or, more frequently, MC simulations.

We also recall that using separate discretizations of stochastic factors is not the only alternative for approximately computing the reliability parameter in stress–strength models or more generally to approximate the distribution of a complex function of stochastic factors. A natural alternative is the direct simultaneous discretization of the multivariate random vector consisting of both stress and strength random factors; this approach would be much more reasonable when the hypothesis of independence between stochastic factors fails and then one needs to take into account, as much as possible, the stochastic dependence among them. A criterion in order to partition the continuous multidimensional support consists in minimizing the loss of mutual information rising from the data reduction process (Colin et al., 2013); an exact dynamic programming (DP) algorithm and an iterative heuristic procedure to perform discretization of an arbitrary n -dimensional pdf are proposed in Christofides et al. (1999).

In this work, we will limit our attention to univariate discretization. In the next section, we will discuss the most used procedures available in the literature.

3 Discretization of a continuous random variable

It is a well-known fact that a large number of machine learning, classification and statistical techniques can only be applied to data sets composed entirely of discrete variables (Dougherty et al., 1995). However, a very large proportion of real data sets include continuous variables; one solution to this issue is to partition continuous variables into a number of sub-ranges and treat each such sub-range as a category, thus simplifying probability calculations and allowing conditional distributions to be assessed more easily. This process of partitioning continuous variables into categories is usually termed “discretization”. We will refer to discretization/approximation also when, for the same reason, a continuous random distribution is substituted/approximated by a discrete random distribution.

Discretization techniques are used in many problems of practical interest in different fields such as statistics, finance, insurance, engineering, etc.

An important application to finance is the development of discrete-time models for pricing options and other derivatives. In this context, a persistent problem arises concerning the accuracy of representation of the financial dynamics, and in particular concerning the discretized approximation of the implied pdf of asset prices (equities, bonds, etc.) (Christofides et al., 1999).

In stochastic programming, the evaluation of the expected value of the continuous stochastic component appearing in the objective function may require, among all, multivariate numerical integration, which is a rather cumbersome task; this obstacle can be overcome by approximating the continuous distribution via a finite number of points, called “scenarios”, which allow to model the stochastic problem as a deterministic optimization problem (see e.g. Ruszczyński and Shapiro, 2003; Šmíd, 2009).

In Tancrez et al. (2011), in the context of analysis of manufacturing systems with finite capacity, the authors suggest discretization of continuous processing time distributions with

finite support into discrete probability distributions with a finite number of points, which result more tractable for subsequent analytical modeling through a Markov chain.

Moreover, there exist many engineering non-linear transformations where the discretization techniques can be used. In this paper, we deal with a well-known engineering problem, formulated as a stress–strength model.

Generally, discretization can be defined as the process that divides continuous values into a finite or at most countable number of intervals, where each interval can be expressed as a value or a symbol (Zhao et al., 2012). Let X be a continuous attribute observed on a sample, and let its range be $S = (l, u) \subseteq \mathbb{R}$. Discretization of the continuous attribute consists of obtaining a “suitable” partition $P = \{[h_0, h_1], [h_1, h_2], \dots, [h_{k-1}, h_k]\}$, in which $l = h_0 < h_1 < \dots < h_{k-1} < h_k = u$; $h_i, i = 1, \dots, k - 1$, are the “breakpoints”. The partition P thus separates the continuous attribute X into k intervals that assure $S = [h_0, h_1] \cup [h_1, h_2] \cup \dots \cup [h_{k-1}, h_k]$; the value of X is replaced by \hat{x} , chosen from the set $S_d = \{\hat{x}_1, \dots, \hat{x}_k\}$, such that if $h_{i-1} \leq X \leq h_i$, then $\hat{x} = \hat{x}_i$, with each \hat{x}_i possibly respecting the consistent constraint $h_{i-1} \leq \hat{x}_i \leq h_i$.

If X is a continuous rv with pdf $f(x)$ defined on S , discretization consists of providing a discrete variable X_d , defined on a support $S_d = \{x_1, \dots, x_k\} \subset S$ (with $x_1 < \dots < x_k$), with probability mass function (pmf) $p_i = P(X_d = x_i), i = 1, \dots, k$, which resembles X according to some criterion.

How to discretize a continuous rv in practice? Obviously, there does not exist a unique answer. Different approaches can be followed, according to which features of the continuous distribution one intends to preserve or which “distance” to the original distribution one intends to minimize. In this sense, a rough classification of the techniques employed for discretizing continuous rv will be outlined in the following subsections. Such a review cannot be exhaustive, since many discretization techniques are scattered through the statistical and information science literature (see, e.g., Hammond and Bickel, 2011, for a partial review); however, it is an attempt to put order and find distinguishing features among the numerous proposals.

3.1 Moment matching

Miller and Rice (1983) stated that “*The accuracy of a discrete approximation of a probability distribution can be measured by the extent to which the moments of the approximation match those of the original distribution*”. Although this assertion is susceptible of criticism, the most popular and long-established approach in finding an approximation of a continuous distribution has been to match as many as possible of its first statistical moments: raw moments (moments about zero), such as the mean, or central moments, such as the variance, or a combination of these can be used.

3.1.1 Gaussian quadrature

A discrete random distribution with k points has $2k - 1$ independent parameters: k values (x_1, \dots, x_k) and $k - 1$ probabilities (say p_1, \dots, p_{k-1}), since the last probability p_k is such that all the p_i sum up to 1. Therefore, a k -point distribution can match $2k - 1$ moments of a given distribution. For example, a 2-point distribution can match the mean, variance and skewness of a given distribution. For any symmetrical distribution, a 3-point approximation yields the following results:

$$\begin{cases} x_1 = \mu - \sqrt{\kappa}\sigma & x_2 = \mu & x_3 = \mu + \sqrt{\kappa}\sigma \\ p_1 = 1/(2\kappa) & p_2 = 1 - 1/\kappa & p_3 = 1/(2\kappa) \end{cases},$$

with $\mu = \mathbb{E}(X)$ and $\kappa = \mathbb{E}((X - \mu)^4)/\sigma^4$. For a normal distribution, for which $\kappa = 3$, these results turn into:

$$\begin{cases} x_1 = \mu - \sqrt{3}\sigma & x_2 = \mu & x_3 = \mu + \sqrt{3}\sigma \\ p_1 = 1/6 & p_2 = 2/3 & p_3 = 1/6 \end{cases}$$

(see, e.g., D’Errico and Zaino, 1988).

In the general case, calculating a k -point discrete distribution matching as many moments as possible of the parent continuous rv X requires solving $2k$ simultaneous non-linear equations for the values x_i and probabilities p_i :

$$\mathbb{E}(X^r) = \sum_{i=1}^k p_i x_i^r, \quad r = 0, 1, \dots, 2k - 1, \tag{4}$$

where the left member is the r -th non-central moment of the continuous rv and the right member is the same order moment of the discretized rv (for $r = 0$, we obtain the condition that the probabilities p_i sum up to 1). System (4) can be solved by resorting to Golub and Welsch method (Golub and Welsch, 1969) for Gaussian quadrature, which reduces to solving for the eigenvalues/eigenvectors of a sparse matrix (see also Toda (2020) for a detailed description of the method). It can be shown that if the original moments of the continuous rv are finite, the method always yields k real, distinct values x_i , which all lie in the interval spanned by the original rv; it also produces positive probabilities p_i . We will refer to the discretization method satisfying system (4) simply as “Gaussian quadrature”, but we recall that the Gaussian quadrature method is more generally an approximate method of calculation of a definite integral; more importantly, this discretization technique, despite its somewhat misleading name, can be applied to both Gaussian and non-Gaussian rv’s.

3.1.2 A two-moment matching technique

Drezner and Zerom (2016) propose a “*simple and effective discretization technique*” for continuous rvs. They consider a continuous rv X with pdf $f(x)$ defined over an interval (a, b) , which is discretized according to a set of thresholds $a = \delta_0 < \delta_1 < \dots < \delta_k = b$. To each sub-interval $(\delta_{i-1}, \delta_i), i = 1, \dots, k$, the probability $p_i = \int_{\delta_{i-1}}^{\delta_i} f(t)dt$ is naturally associated; a representative value of the interval is the conditional first moment $m_i = \frac{1}{p_i} \int_{\delta_{i-1}}^{\delta_i} t f(t)dt$. The rv $M = \{(m_i, p_i), i = 1, \dots, k\}$ thus arises as a possible discrete counterpart to X . It can be proved that for any choice of the thresholds $\{\delta_i, i = 1, \dots, k - 1\}$, $\mu_M = \mu_X$, whereas $\sigma_M^2 < \sigma_X^2$. This means that M preserves the expected value of the underlying continuous distribution, but cannot preserve its variance, which is underestimated; the idea of the authors is to find a discrete distribution which is “close” to M but is also able to preserve the variance of X . With this aim, a discrete rv $D = \{(d_i, p_i); i = 1, \dots, k\}$ is obtained as the solution to the following optimization problem:

$$\left\{ \begin{array}{l} \min_{p_i, d_i; i=1, \dots, k} \sum_{i=1}^k p_i (m_i - d_i)^2 \\ \text{subject to} \\ \sum_{i=1}^k p_i = 1 \\ \sum_{i=1}^k p_i d_i = \mu_X \text{ (preservation of the first moment)} \\ \sum_{i=1}^k p_i d_i^2 = \mu_X^2 + \sigma_X^2 \text{ (preservation of the variance)} \end{array} \right. \quad (5)$$

with $p_i = P(M = m_i) = P(D = d_i) = \int_{\delta_{i-1}}^{\delta_i} f(t) dt$; $m_i = \frac{1}{p_i} \int_{\delta_{i-1}}^{\delta_i} t f(t) dt$.

According to (5), D is the discrete rv which is the closest in squared error sense to $M = \{(m_i, p_i), i = 1, \dots, k\}$ and which preserves the first two moments of X (see also Fig. 1 for a graphical explanation).

The non-linear optimization problem stated above can be further simplified and proved to be equivalent to the following one:

$$\left\{ \begin{array}{l} \max_{p_i, i=1, \dots, k} \sum_{i=1}^k p_i m_i^2 \\ \text{subject to} \\ \sum_{i=1}^k p_i = 1 \\ d_i = (m_i - \mu_X) \sigma_X / \sigma_M + \mu_X, \quad i = 1, \dots, k \end{array} \right.$$

This latter maximization problem has been implemented by the authors under the R environment (R Core Team, 2019) for the standard normal and a Weibull distribution, but can be easily customized to any continuous cdf whose expressions for the first two moments are available. Although the discrete approximation is provided numerically, this technique has a wider applicability than the Gaussian quadrature technique described in Sect. 3.1.1, since it requires the existence of the first two moments only of the continuous rv.

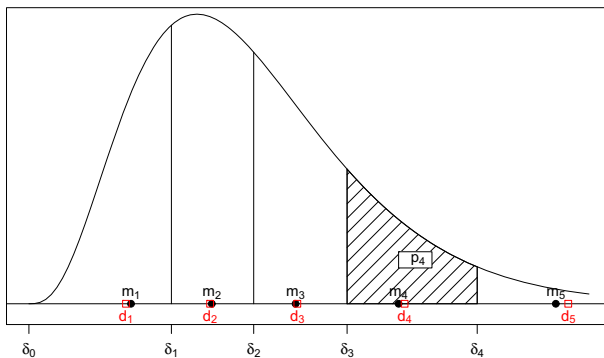


Fig. 1 Discretization of a continuous probability distribution by $k = 5$ points according to the method suggested by Drezner and Zerom (2016). The graph displays the curve of the pdf of a continuous rv X , which has to be approximated by a discrete rv D , by using an intermediate discrete rv M (see Sect. 3.1.2 for details). Each value m_i (represented through filled black circles) is computed as the mean of the segment (δ_{i-1}, δ_i) with probability p_i equal to the area under the curve of the pdf for that segment; the resulting rv M preserves the mean but not the variance of X ($\sigma_M^2 < \sigma_X^2$ for any choice of $\{\delta_i; i = 0, \dots, k\}$). The approximating discrete distribution D (whose values are displayed in the graph through empty red squares), is the closest in squared error sense to M , sharing its probabilities p_i , while preserving the first two moments of X

3.2 Preservation of the survival or distribution function

3.2.1 Matching the distributions at k points

Since moments are only partial measures of the distribution form and discretization based on moment equalization up to a finite order cannot retain the functional properties of the original distribution, another appealing approach is to match the cdf, or alternatively, the survival function, of the continuous and discrete rvs at the k support points of the latter. A first proposal was suggested by Roy and Dasgupta (2001). Letting F_x be the cdf of the continuous rv X , if the discretized rv X_d takes only integer values, its pmf can be assigned as

$$P(X_d = i) = F_x(i + \delta) - F_x(i - (1 - \delta)) \quad i \in \mathbb{Z}, 0 < \delta < 1. \tag{6}$$

Except for a shift in the location by δ , Eq. (6) retains the form of the original cdf. To reduce the range of X_d , a standard form of Eq. (6) can be first considered. For example, if $X' \sim N(\mu, \sigma^2)$, the discretized form of X' is $X'_d = \mu + \sigma \cdot X_d$, where X_d is the discretized version of $X \sim N(0, 1)$. For the k -point discretization, k odd, X_d takes the integer values from $x_1 = -(k - 1)/2$ to $x_k = (k - 1)/2$, and the probabilities for the minimum and maximum values are determined as $P(X_d = x_1) = F_x(x_1 + \delta)$ and $P(X_d = x_k) = 1 - F_x(x_k - (1 - \delta))$. As for the value of δ in Eq. (6), its optimum choice should be such that the proposed discrete pmf can retain the first (and/or second) raw moments of X . For symmetrical unimodal distributions the optimal value is 0.5.

Roy and Dasgupta (2002) and Roy (2004) extended the discretization approach in Roy and Dasgupta (2001) to two asymmetrical distributions defined on \mathbb{R}^+ , the Weibull and Rayleigh distributions respectively, after defining the parameter values leading to a “standard” version of the distribution.

A more general heuristic approach, applicable to any continuous distribution, is proposed in Barbiero (2012), and is still based on Eq. (6). The “standard” interval $(-g, +g)$, with $g > 0$ a parameter fixed “a priori”, is chosen and is divided into $k - 1$ equally spaced sub-intervals whose extremes are k equally-spaced “standard” points z_i :

$$z_i = -g + \frac{2g(i - 1)}{k - 1} \quad i = 1, \dots, k.$$

The k points for the discretized version of a general rv X with cdf $F_x(x)$ are determined in the following way. The cumulative probabilities of z_i are calculated as $u_i = \Phi(z_i)$. Then, the points of the support of the discretized rv X_d are obtained by applying the inverse cdf of X , as $x_i = F_x^{-1}(u_i)$. The $k - 1$ interval central values $x_i^* = (x_i + x_{i+1})/2, i = 1, \dots, k - 1$ are considered, and the probabilities of the points x_i are finally provided by

$$\begin{cases} p_1 = P(X_d = x_1) = F_x(x_1^*) \\ p_i = P(X_d = x_i) = F_x(x_i^*) - F_x(x_{i-1}^*) \quad i = 2, \dots, k - 1 \\ p_k = P(X_d = x_k) = 1 - F_x(x_{k-1}^*). \end{cases}$$

This procedure does not retain the expected value of the original continuous variable (unless it is symmetrical), nor its higher order moments, yet it tries to retain its characteristics in terms of the cdf. The approach can be easily applied to any continuous distribution and its simplicity makes it suitable for software automation and implementation. The procedure is also flexible, since it allows the user to choose, along with the number k of points for the discretization, the additional parameter g for setting up the interval for the standard normal and finally the

range of its discrete approximation. On the other hand, this poses an additional problem, which value of g to choose. Some indications are reported in Barbiero (2012), suggesting that g should take values between 2 and 3 if the number of point k lies between 5 and 9.

3.2.2 Minimizing a distance between the two distributions

Another way of finding a discrete counterpart of the continuous random distribution is to minimize some proper distance between the original cdf and the cdf of the approximating discrete rv. Let us suppose that F_x is a strictly increasing cdf defined on the real line, and \hat{F} is an approximation of this distribution. If we define the distance between F and \hat{F} using the L^r metric ($r > 0$)

$$\|F_x - \hat{F}\|_r = \int_{\mathbb{R}} |F_x(x) - \hat{F}(x)|^r dF(x) = \int_0^1 |t - \hat{F}(F_x^{-1}(t))|^r dt \quad (7)$$

then the best discrete k -point approximation - i.e., the approximation that yields the minimum value of this distance - has been proved (Kennan, 2006) to have equally-weighted support points x_i (i.e., with probability $1/k$ each) such that $F_x(x_i) = \frac{2i-1}{2k}$ or, equivalently,

$$x_i = F_x^{-1}\left(\frac{2i-1}{2k}\right), \quad (8)$$

the minimal distance between F_x and \hat{F} is

$$\|F_x - \hat{F}\|_r = \frac{(2k)^{-r}}{r+1}. \quad (9)$$

It is curious that this approximation method coincides with the so called “bracket-median” method introduced by Clemen (1991) and taken back by Smith (1993), even if its features were probably not fully clear and appreciated. It is noteworthy that this best approximate distribution is always a discrete uniform distribution and its support points do not depend on which L^r norm is used. This is somewhat surprising and distinguishes Kennan’s technique from all the other techniques presented so far. Kennan’s approach looks very attractive because once the number of points is chosen, their value and their probabilities are automatically and simultaneously derived in a straightforward manner, which can be applied to any continuous distribution (possessing finite moments or not). Note that the cdf of the discrete rv is

$$\hat{F}(x) = \begin{cases} 0 & x < F_x^{-1}(1/(2k)) \\ i/k & F_x^{-1}((2i-1)/(2k)) \leq x < F_x^{-1}((2i+1)/(2k)), i = 1, \dots, k-1 \\ 1 & x \geq F_x^{-1}((2k-1)/(2k)) \end{cases}$$

so that $\forall x \in \mathbb{R} |F_x(x) - \hat{F}(x)| \leq 1/(2k)$, which tends to zero for $k \rightarrow \infty$. This means that $\lim_{k \rightarrow \infty} \hat{F}(x) = F_x(x) \forall x \in \mathbb{R}$, i.e., the approximating discrete rv X_d converges in distribution to X .

3.3 Hybrid techniques

Many other discretization techniques have been proposed in the literature, which do not fall in either broad class illustrated so far. For example, Ghosh and Roy (2011) proposed an hybrid approximation technique introducing the concept of linear transformation of the discretized variable for equalization of first moments of a continuous variable, which, in fact,

strikes a balance between the moment equalization techniques and discrete concentration techniques. Here the values x_i are fixed in advance (for example, for a standard normal to be approximated by 7 points one can choose $-3, -2, \dots, +2, +3$) and the values $p_i = P(X_d = x_i)$ are determined under a linear set-up under the condition of equality of raw moments, $\mathbf{A}\mathbf{p} = \boldsymbol{\mu}$, where $\mathbf{p} = (p_1 \dots p_k)^T$ is the vector of probabilities, $\boldsymbol{\mu} = (\mu_0 \mu_1 \dots \mu_{k-1})^T$ is the raw moment vector of X with $\mu_0 = 1, \mu_k = \mathbb{E}(X^k), k \geq 1$, and the $k \times k$ matrix $\mathbf{A} = (a_{ij}) = (x_j^{i-1})$ is constructed out of prefixed k points, which makes the problem indeed simpler to solve. The analytical solution to the previous equation is obviously $\mathbf{p} = \mathbf{A}^{-1}\boldsymbol{\mu}$, which, however, is not guaranteed to be feasible (i.e., some p_i may not satisfy the essential condition to lay in the unit interval).

Tanaka and Toda (2013) proposed a method that minimizes the Kullback-Leibler distance (Kullback, 1959) between the target discrete distribution and a discrete prior distribution, under some constraints on the moments. In more detail, once the support $S_d = \{x_1, \dots, x_k\}$ is chosen, the optimization problem is formulated as:

$$\left\{ \begin{array}{l} \min_{p_i, i=1, \dots, k} \sum_{i=1}^k p_i \log(p_i/q_i) \\ \text{subject to} \\ \mathbf{A}\mathbf{p} = \boldsymbol{\mu} \end{array} \right.$$

where $q_i = f(x_i) / \sum_{i=1}^k f(x_i)$ defines a discrete prior distribution, and $\mathbf{p}, \mathbf{A}, \boldsymbol{\mu}$ have the same meaning as before, with the exception that \mathbf{A} is a $l \times k$ matrix, with $l \leq k$. The problem reduces to an unconstrained convex optimization problem and then global convergence is guaranteed by algorithms such as the Newton-Raphson or the steepest descent methods. We highlight that in this method the problem of finding the discrete points is solved separately from the problem of assigning the probabilities. In an example, the authors consider the standard normal distribution and choose equally spaced points on $[-4, 4]$ (with distance 0.1); the same choice is made for a beta distribution.

3.4 An example

In this subsection, we consider two continuous random distributions, a standard normal and a Weibull, and show how the discretization techniques presented so far practically work. Table 1 reports the probability mass functions of the discrete rvs with $k = 9$ points derived through the Gaussian quadrature (abbreviated as GQ), the discretization procedures proposed by Roy and Dasgupta (2001, 2002) (RD), Barbiero (2012) (B), with $g = 2.5$, Kennan (2006) (K), and Drezner and Zerom (2016) (DZ), from a standard normal and a “standard” Weibull distribution. This latter is obtained as a nearly-symmetric small range Weibull distribution, with pdf $f(x; \lambda, \beta) = \frac{\beta}{\lambda} \left(\frac{x}{\lambda}\right)^{\beta-1} e^{-(x/\lambda)^\beta}, x > 0$, by setting $\lambda = 5$ and $\beta = 3.5$ (Roy and Dasgupta, 2002). One can easily appreciate the differences by moving through the discrete distributions in terms of both support values x_i and probabilities p_i .

Note in particular the differences between the classical quadrature method (preserving up to the first $2k - 1$ moments) and the method in Drezner and Zerom (2016) (preserving only the first two moments); the former produces much more extreme values for x_1 and x_k with much smaller associated probabilities (equal to zero at the fourth decimal digit).

We remark that for both situations, the method by Kennan (2006) yields the smallest range for the discrete approximation, whereas the Gaussian quadrature yields the largest one.

Table 1 Probability mass function for a continuous rv discretized according to Gaussian Quadrature (GQ) Roy and Dasgupta (2001, 2002), Roy (2004) (RD), Barbiero (2012) (B), Kenman (2006) (K), Drezner and Zerom (2016) (DZ) discretization methods (number of points $k = 9$)

method	GQ			RD			B ($g = 2.5$)			K			DZ		
	x_i	P_i		x_i	P_i		x_i	P_i		x_i	P_i		x_i	P_i	
<i>(a) Standard normal</i>															
1	-4.513	0.0000		-4	0.0002		-2.5	0.0144		-1.593	0.1111		-2.287	0.0311	
2	-3.205	0.0028		-3	0.0060		-1.875	0.0447		-0.967	0.1111		-1.497	0.0845	
3	-2.077	0.0499		-2	0.0606		-1.25	0.1152		-0.589	0.1111		-0.932	0.1323	
4	-1.023	0.2441		-1	0.2417		-0.625	0.2031		-0.282	0.1111		-0.450	0.1644	
5	0	0.4063		0	0.3829		0	0.2453		0	0.1111		0	0.1755	
6	1.023	0.2441		1	0.2417		0.625	0.2031		0.282	0.1111		0.450	0.1644	
7	2.077	0.0499		2	0.0606		1.25	0.1152		0.589	0.1111		0.932	0.1323	
8	3.205	0.0028		3	0.0060		1.875	0.0447		0.967	0.1111		1.497	0.0845	
9	4.513	0.0000		4	0.0002		2.5	0.0144		1.593	0.1111		2.287	0.0311	
<i>(b) "Standard" Weibull, with pdf $f(x; \lambda, \beta) = \frac{\beta}{\lambda} (\frac{x}{\lambda})^{\beta-1} e^{-(x/\lambda)^\beta}$, $x > 0, \lambda = 5, \beta = 3.5$, see Roy and Dasgupta (2002)</i>															
1	0.616	0.0027		0.5	0.0036		1.172	0.0151		2.207	0.1111		1.577	0.0454	
2	1.470	0.0317		1.5	0.0361		1.851	0.0453		3.075	0.1111		2.523	0.0949	
3	2.544	0.1355		2.5	0.1144		2.673	0.1150		3.628	0.1111		3.266	0.1326	
4	3.744	0.2918		3.5	0.2133		3.575	0.2020		4.084	0.1111		3.924	0.1567	
5	4.996	0.3251		4.5	0.2647		4.503	0.2445		4.503	0.1111		4.548	0.1650	
6	6.260	0.1728		5.5	0.2173		5.418	0.2031		4.919	0.1111		5.173	0.1560	
7	7.528	0.0377		6.5	0.1117		6.302	0.1156		5.367	0.1111		5.835	0.1291	
8	8.829	0.0026		7.5	0.0333		7.148	0.0450		5.907	0.1111		6.598	0.0863	
9	10.253	0.0000		8.5	0.0056		7.956	0.0145		6.771	0.1111		7.630	0.0340	

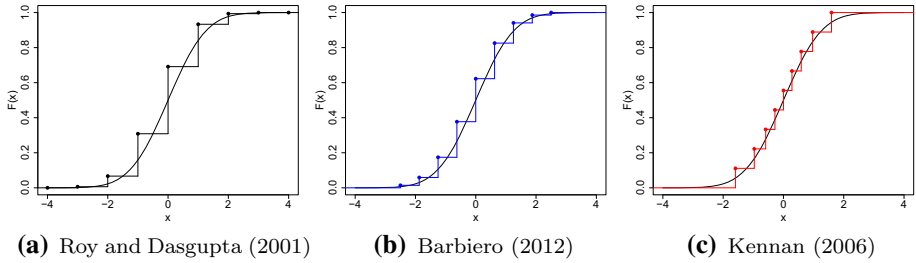


Fig. 2 Cumulative distribution function of a standard normal rv (solid line) and its discrete 9-point approximation (step-wise curve) derived according to three techniques

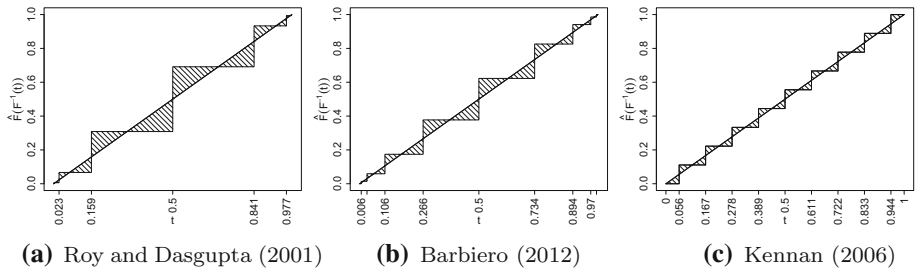


Fig. 3 Graphs of $\hat{F}(F^{-1}(t))$ and distances $\|F - \hat{F}\|_1$ (shaded areas) for a standard normal and its discrete approximations according to three techniques ($k = 9$). Note that in a) only 5 points out of 9 have been drawn for the sake of visibility

It is also worth noting that for the standard normal distribution, all the discrete approximations maintain the symmetry of their parent distribution and then the expected value is equal to 0. As for the variance, Kennan’s equal-probability discrete distribution largely underestimates the original variance, whereas Roy and Dasgupta (2001) and Barbiero (2012) slightly overestimate it: infact, for RD, $\text{Var}(X_d) = 1.083$, for B, $\text{Var}(X_d) = 1.012$, and for K, $\text{Var}(X_d) = 0.867$. The Gaussian Quadrature method and the discretization technique by Drezner and Zerom (2016), as we know, preserve the unit variance.

The difference among the discretization techniques is evident also looking at Fig. 2, where the curve of the cdf of the standard normal is displayed, along with the graphs of the step functions representing the cdf of some of its discretized versions. From Fig. 3, it clearly emerges that, as expected, the distance between the continuous cdf and the cdfs of discrete approximations, in this case, for Kennan’s procedure, by using (9), $\|F - \hat{F}\|_1 = 1/36$.

3.5 A general statistical application

We now introduce and discuss a possible application of discretization to a simple but notable statistical problem. We want to determine the p -quantile ($0 < p < 1$) of the distribution of the sum of n independent continuous rvs. This is a widely investigated problem, also in quantitative finance: if each single rv represents a risk related to a single asset, then the sum can be regarded as the aggregate risk of the portfolio comprising the n assets. If such risks can be interpreted as loss distributions for a given time horizon, then computing the p -quantile of the sum means computing the corresponding Value-at-Risk at level p for the whole portfolio.

Table 2 True and approximate values of the p -quantile for the sum of three Gamma rvs ($\alpha_1 = 3, \alpha_2 = 6, \alpha_3 = 9, \beta = 2$)

p -quantile	90%	95%	97.50%	99%
TRUE	11.80	12.75	13.61	14.65
NORM	11.72	12.49	13.16	13.93
GQ	11.85	12.92	13.53	14.90
B	11.75	12.79	13.38	14.60
K	<i>11.48</i>	<i>12.20</i>	<i>12.82</i>	<i>13.46</i>
D	11.83	12.69	13.63	14.60

NORM, NORMal approximation; GQ, (Discretization by) Gaussian Quadrature; B, (Discretization by) Barbiero (2012); K, (Discretization by) Kennan (2006); DZ, (Discretization by) Drezner and Zerom (2016)

In order to recover the p -quantile of a continuous rv, we need to be able to solve the equation $F(q) = p$ in terms of q ; this is possible if the expression of the cdf F of the sum is known analytically. For example, we know that if $X_1, \dots, X_i, \dots, X_n$ are n independent Gamma rvs with parameters (α_i, β) (shape and rate parameters, respectively), $i = 1, \dots, n$, then the sum $S = \sum_{i=1}^n X_i$ is still a Gamma rv with parameters $\alpha = \sum_{i=1}^n \alpha_i$ and β . Then its cdf is analytically known and the p -quantile can be easily determined. For other parametric families of distributions, finding the cdf of the sum is not so straightforward. If each rv X_i has finite first moment and variance, μ_i and σ_i^2 , then we know that $\mu_S = \sum_{i=1}^n \mu_i$ and $\sigma_S^2 = \sum_{i=1}^n \sigma_i^2$ and we can approximate the distribution of S through the Gaussian law, by resorting to a relaxed version of the central limit theorem (Lyapunov condition): thus, the p -quantile of S can be approximated by $\mu_S + z_p \sigma_S$. Alternatively, one can consider discretizing the n rvs through k points each, and construct a discretized version of the sum S based on a grid of k^n points and associated probabilities. Then, the probability mass function of the discretized S can be reckoned and its corresponding p -quantile can be computed, as the minimum value for which the cumulative probability is greater than or equal to p . For the case of n gamma rvs, along with the exact value of the quantile, two approximated values can be computed: one from normal approximation (recalling that $\mu_i = \alpha_i \beta$ and $\sigma_i^2 = \alpha_i \beta^2$), the other from discretization. Clearly, since in this case we are able to compute the exact quantiles easily, we would not need to resort to any approximation, but we do it just to assess the accuracy of approximation. Table 2 reports the results for the computation of the p -quantile of a sum, when we consider $n = 3$ independent Gamma rvs with parameters $\alpha_1 = 3, \alpha_2 = 6, \alpha_3 = 9, \beta = 2$, with $p = 0.9; 0.95; 0.975; 0.99$. The first row displays the “true values” of quantiles; the other rows display the values obtained through the normal approximation or through different discretization techniques and enumeration. For each value of p , figures in bold font are the closest ones to the corresponding true value; figures in italic font are the farthest ones. As one can see, the discretization method proposed by Drezner and Zerom (2016) provides the best approximation in three cases out of four; the one proposed by Kennan (2006) yields the worse approximation in all the cases.

In the next section, we will discuss an application in the context of stress–strength models through a comparative closeness study, where we empirically evaluate the approximation-by-discretization techniques here introduced and the FORM method recalled in the previous section.

4 A comparative closeness study

In this section, the main results of a comparative study will be presented, which has been carried out with the aim of evaluating the performance of the discretization methods described in Sect. 3 with respect to the practical problem outlined in Sect. 2. Specifically, we will consider the Gaussian quadrature method and the methods proposed by Roy and Dasgupta (2001, 2002), Barbiero (2012), Kennan (2006), and Drezner and Zerom (2016). The key question we will deal with is whether a discretization method turns out to be superior to others - i.e., if it performs uniformly better - or, if this is not the case, under which conditions a particular method can outperform others. Moreover, we will investigate the effect of the number of points on the statistical performance of the discretization methods: obviously, we expect that increasing the number of discrete points would enhance it.

A well-known engineering stress–strength problem will be considered, where the reliability parameter can be either “exactly” computed recalling MC simulations via Eq. (2) or approximately assessed according to FORM, see Sect. 2, or through the approximation-by-discretization approach synthesized by Eq. (3), which uses one of the techniques described in Sect. 3. The accuracy of the approximation methods in recovering the “true” value of reliability, as provided by MC simulations, as well as their computational burden will be evaluated.

4.1 Simulation design

The design of this comparative study is very similar to that carried out in Barbiero (2012), which reprised those in Roy and Dasgupta (2001, 2002), with some modifications that tend to make the results more reliable. We will focus on a context where the rv Y representing stress is a function of several random sub-factors, while rv X , representing strength, does not depend on random sub-factors, but is a rv itself. X is assumed to have a fixed variance, while its expected value varies within an interval by a constant step size; all the stochastic stress factors have their distributional parameters, or first moments, fixed to some prespecified values. We will first consider the case where all the stochastic factors involved are assumed to follow the normal distribution, and then we will move to the case where they all follow the Weibull distribution. As highlighted in Roy and Dasgupta (2002), this latter distribution can describe engineering items more appropriately, “*as it provides an excellent way to focus on both burn-in and burnout phenomena and can model increasing failure rate (IFR) and decreasing failure rate (DFR) class of distributions for different choices of the shape parameter*”. For comparative purpose, each homologous stochastic factor shares the same expected value and variance in the two cases. Applying the approaches described in the previous sections, under each experimental condition for the stress–strength model, we compute an approximate value of $R = P(X > Y)$, according to Eq. (3): we consider (i) the Gaussian quadrature technique (abbreviated as GQ), (ii) the proposal in Roy and Dasgupta (2001) and Roy and Dasgupta (2002) (RD), for the normal and Weibull case respectively, (iii) the proposal in Barbiero (2012) (B), (iv) the proposal in Kennan (2006) (K), (v) the proposal in Drezner and Zerom (2016) (DZ). The number of discrete points k varies from 5 to 11, which are the values usually encountered in the applications. In particular, the value $k = 9$ has been used in Roy and Dasgupta (2002) and Barbiero (2012), and seems suitable to find a compromise between accuracy of the approximation and computational efficiency. An approximate value of R is computed through FORM as well. As a true (simulated) value of R , we consider the value provided by (2), obtained through $S = 10,000,000$ MC simulation runs.

Since the objective of discretization of a continuous distribution is to gain in computational efficiency, we will not consider in this simulation study the methods described in Sect. 3.3, which either introduce further complex calculations or may not be able to provide a feasible solution under each simulated experimental condition.

An index for evaluating the performance of each method under a specific experimental condition (here, for each examined value of the expectation μ_X of the strength factor, being all the parameters of the stress factor Y fixed) is the simple difference (“Deviation” or “Bias”) between the approximated/estimated value of R provided by the method and its “true” value: $\hat{R} - R$ (according to the method considered, \hat{R} is substituted by one of \hat{R}_{GQ} , \hat{R}_{RD} , \hat{R}_B , \hat{R}_K , \hat{R}_{DZ} , or \hat{R}_{FORM}). Leaving the sign apart, one can consider (Roy and Dasgupta, 2001) its absolute value, which is denoted as Absolute Deviation (AD): $AD = |\hat{R} - R|$ and is an easy and understandable measure of the closeness of the method’s estimate to the true value. As an overall performance index, following again Roy and Dasgupta (2001), we refer to the Mean Absolute Deviation (MAD), which is the simple average of the absolute deviations computed over all the simulated scenarios.

Obviously, there is some arbitrariness in defining such index, since it is not clear how one should choose the scenarios to simulate on which to compute the average value. Here, we consider equally spaced values of the expected value μ_X of the strength rv, varying from a minimum to a maximum value, corresponding respectively to a value of the reliability parameter almost equal to zero (0.01) and almost equal to unity (0.99) when all the stochastic factors follow a normal distribution. The step on the X variable is taken as small as possible, in order to make results more reliable and not tied to the particular values of X used; ideally, X should range over through a continuum of values.

As a possible future perspective, we could think of modeling the expectation μ_X of the strength X as a rv itself, as happens in the Bayesian context; in this way, we would be able to define theoretically the MAD of an estimator \hat{R} (obtained from some discretization procedure) as the integral of $|R(\mu_X) - \hat{R}(\mu_X)|$ weighted by the pdf of μ_X : $\int_M |R(\mu_X) - \hat{R}(\mu_X)| f_{\mu_X}(\mu_X) d\mu_X$, with M the support of μ_X . Clearly, this integral cannot be computed analytically, but it can be approximated at its time over a dense grid of values for μ_X , but now with a different weight derived by the pdf of the rv μ_X .

In order to graphically evaluate and compare the approximation accuracy of the methods, the boxplots of the distributions of the AD can be used, as well as the dispersion plots of the simulated values of R versus the corresponding values \hat{R} computed for the considered approximation techniques in each scenario: the stronger the linear relationship, measured in this study by Pearson’s correlation coefficient, the better the approximation.

In the next subsection, we will illustrate and comment the results of the comparative closeness study.

4.2 Application: hollow cylinder

The maximum shear stress that a shaft with a hollow cylindrical section can transmit is given by:

$$Y = \frac{16Ma}{\pi(a^4 - b^4)}, \quad (10)$$

where a and b are the outer and inner diameters, and M the applied torque. The strength of the shaft is denoted as X .

4.2.1 Normal case

Let us suppose a , b , and M are rvs following a normal distribution with expected value 2.4, 2 and 1, 200, and standard deviation 0.02, 0.02, 60 respectively. Let us assume that X is a normal rv too, whose expected value varies from 660 to 1100, with unit steps, while its standard deviation is set equal to 55.

With this choice of distributions' parameters, the closeness study has been carried out. Figure 4 displays the dispersion plots between true and estimated values of the reliability parameter derived by the five discretization methods when $k = 9$, for each integer value of μ_X in [660, 1100]. Note that all the plots indicate that the function relating the estimate with the simulated value is very close to the identity function; although with some visible irregularities for GQ and RD; furthermore, for the method by Kennan (2006), the graph points out a slight departure from the first-third quadrant bisector, since we can notice that values of R smaller than 0.5 are underestimated (the graph lies under the bisector), whereas values of R larger than 0.5 are overestimated (the graph lies over the bisector). Summarizing, the correlation coefficients between the MC value of R and its approximation is 0.9999019 for GQ, 0.9998804 (the smallest value) for RD, 0.9999879 for B, 0.9999502 for K, and 0.9999962 (the largest value) for DZ.

By computing the absolute deviations $|R - \hat{R}|$, we notice that the number of times the method of Drezner and Zerom (2016) yielded the minimum value of the absolute error is 279 out of 441; it was 102 for B, 36 for GQ, 19 for RD, and 5 for K. Figure 5 displays the boxplot of the empirical distribution of the AD for the five methods explored. It is clear that the distribution derived by the DZ method is the most shifted towards smaller values (with AD values all smaller than 0.004), followed, in ascending order of the corresponding median value, by AB, GQ, RD, and K.

A more quantitative synthesis of the results is reported in Table 3, which shows the values of the mean deviation (MD) and MAD for the discretization techniques here investigated for $k \in \{5, 6, 7, 8, 9, 10, 11\}$. In this table, we reported also the values of MD and MAD obtained from FORM. In terms of MAD, FORM is still competitive with respect to the discretization methods (although it is always overcome by DZ for any k , and by GQ for $k \geq 10$ and B for $k \geq 8$); however, for each scenario, it always provides an overestimate of R so that the deviation is always greater than zero and then the MD coincides with MAD. Figure 6 graphically displays the values of MAD. From there, it is easy to evaluate the performances comparatively, and to assess their evolution with respect to k . The overall best performer is the method proposed by Drezner and Zerom (2016); its MAD, for each choice of k , is smaller than that of its competitors, and its MD is almost always the smallest in absolute value. The worst method is that proposed by Kennan (2006), which always presents the largest values for MAD and MD (in absolute value), even if its performance improves very rapidly as k increases. We notice that all the methods except Kennan's show a negative MD (i.e., on average they underestimate R). Most methods present an expected downward trend of MAD with respect to k ; the method by Drezner and Zerom (2016) shows a slowly decreasing trend of MAD; for $k = 10$ and 11, it is nearly reached by the technique of Barbiero (2012). The method proposed by Roy and Dasgupta (2001) represents an exception, since the behaviour of MAD is not monotone (decreasing) at all; this fact can be ascribed to the selection of the k points—see Sect. 3.2.1 and the example in Sect. 3.4—which, as k increases, can include values much far from the distribution center with a very small probability, somehow wasting the increased cardinality.

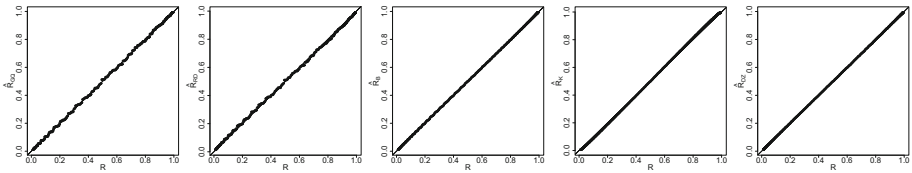


Fig. 4 Dispersion plots between true and estimated reliability parameter (from left to right: GQ, RD, B, K, and DZ method) when $k = 9$ and stochastic sub-factors are normal

Fig. 5 Boxplots of the distribution of Absolute Deviation for the discretization methods (from left to right: GQ, RD, B, K, and DZ method) when $k = 9$ and stochastic sub-factors are normal

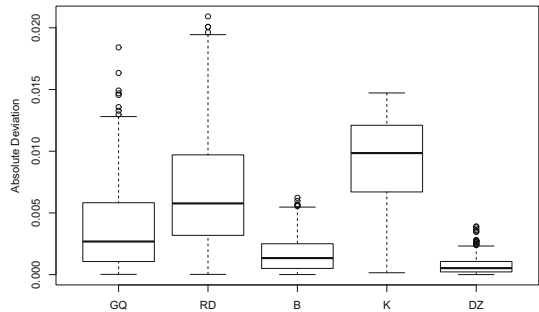


Table 3 Results of the comparative study for the hollow cylinder stress–strength problem with normal distributions; values of MD and MAD (to be multiplied by 10^{-3}) are displayed for FORM and for different discretization techniques and for different values of k

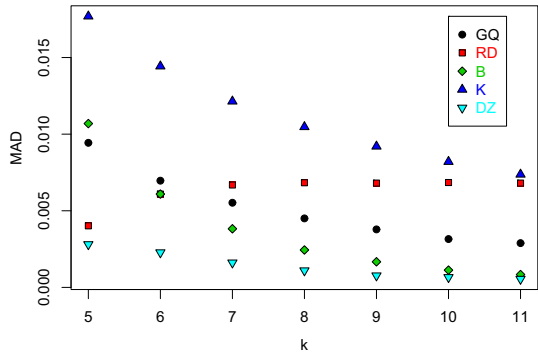
k	Method:	GQ	RD	B	K	DZ	FORM
5	MD	0.075	0.366	1.319	-2.219	0.060	3.444
	MAD	9.434	4.018	10.688	17.695	2.807	3.444
6	MD	0.313	0.784	0.784	-1.873	0.109	3.444
	MAD	6.964	6.090	6.090	14.430	2.277	3.444
7	MD	0.110	0.971	0.490	-1.578	0.078	3.444
	MAD	5.524	6.682	3.824	12.143	1.612	3.444
8	MD	0.137	0.936	0.348	-1.399	0.068	3.444
	MAD	4.503	6.828	2.442	10.467	1.099	3.444
9	MD	0.054	0.999	0.251	-1.220	0.097	3.444
	MAD	3.785	6.806	1.670	9.201	0.769	3.444
10	MD	0.160	0.939	0.163	-1.103	0.137	3.444
	MAD	3.155	6.843	1.126	8.201	0.661	3.444
11	MD	0.079	0.999	0.104	-0.999	0.124	3.444
	MAD	2.889	6.807	0.837	7.371	0.546	3.444

4.2.2 Weibull case

The values of R and its approximate estimates have been then computed under the assumption that M, a, b and X are Weibull rvs with the same values of mean and variance as before. We recall that for a Weibull distribution with pdf

$$f(x; \lambda, \beta) = \frac{\beta}{\lambda} \left(\frac{x}{\lambda}\right)^{\beta-1} e^{-(x/\lambda)^\beta}, \quad x > 0, \lambda, \beta > 0;$$

Fig. 6 Values of MAD for the discretization techniques applied to the stress–strength problem with normal distributions. Each technique is depicted through a different symbol and colour



the r integer moment has the following expression:

$$\mathbb{E}(X^r) = \lambda^r \Gamma(1 + r/\beta), \quad r = 1, 2, \dots,$$

therefore, expected value and variance are equal to

$$\begin{aligned} \mu_X &= \mathbb{E}(X) = \lambda \Gamma(1 + 1/\beta) \\ \text{Var}(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \lambda^2 [\Gamma(1 + 2/\beta) - \Gamma^2(1 + 1/\beta)]. \end{aligned}$$

Thus, it is quite straightforward to derive the values of the parameters λ and β ensuring target values of expectation and variance, by solving the system consisting of the two equations above in terms of the unknown λ and β : although it cannot be solved analytically, an iterative numerical procedure can be implemented for this aim. For example, if we require an expected value equal to 2.4 and variance equal to 0.0004 (these are the values of the stress stochastic sub-factor a in Eq. (10)), then the parameters must be set as $\lambda = 2.408977$ and $\beta = 153.176152$.

We observe that passing from normal to Weibull distribution in the stress–strength problem studied does not alter the relative performance of the discretization methods when recovering the reliability parameter. Yet we emphasize that the Gaussian quadrature method was not able to produce feasible results in our comparative study, although in general it can be applied also to non-Gaussian distributions as the Weibull, see Sect. 3.4. In fact, for the choice of parameters used here for some stochastic factors, characterized by a small value of the standard deviation related to the expected value, the algorithm regulating discretization can produce unexpected (negative) values for the discrete points, or cannot produce a feasible solution at all, even for small values of k . This is related to numerical issues affecting the eigenvalue/eigenvector problem at the basis of Golub and Welsch’ algorithm, which lead to an intermediate matrix that is not positive definite as theoretically expected. Therefore, the Gaussian quadrature was excluded from the comparative study. As for RD technique, we report only the results for $k = 9$, as in Roy and Dasgupta (2002) the discretization procedure for the Weibull distribution is explicitly conceived for that value only. Figure 7 displays the dispersion plots between true and estimated values of the reliability parameter derived by the four discretization methods RD, B, K, and DZ when $k = 9$, for each integer value of μ_X in [660, 1100]. Note that all the plots indicate that the function relating the estimate with the simulated value is very close to the identity function, although with some visible irregularities for RD and B; as in the normal case, for the method by Kennan (2006), the graph lies under the first-third quadrant bisector for values of R smaller than 0.5, whereas it lies above for values of R larger than 0.5. Summarizing, the correlation coefficient between the MC value of R and its approximation

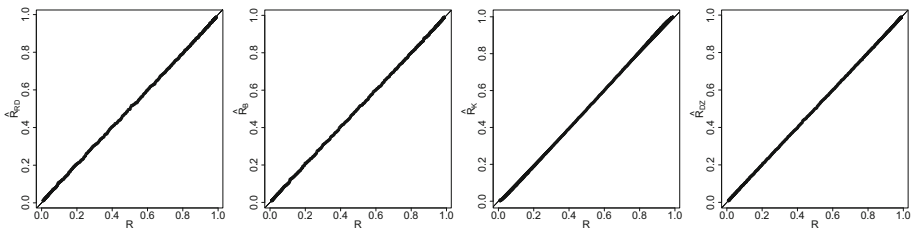
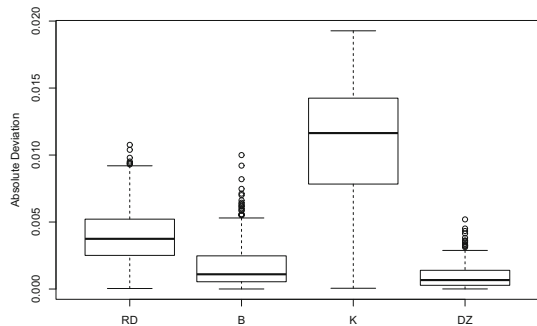


Fig. 7 Dispersion plots between true and estimated reliability parameter (from left to right: RD, B, K, and DZ method) when $k = 9$ and stochastic sub-factors follow the Weibull distribution

Fig. 8 Boxplot of the distribution of Absolute Deviation for the discretization methods (from left to right: RD, B, K, and DZ method) when $k = 9$ and stochastic sub-factors follow the Weibull distribution



is 0.9999755 for RD, 0.9999833 for B, 0.9999584 (the smallest value) for K, and 0.9999955 (the largest value) for DZ.

If we consider the absolute deviations $|R - \hat{R}|$, the number of times the method of Drezner and Zerom (2016) yielded the minimum value of the absolute error is 277 out of 441; it is 137 for B, 18 for RD, and 9 for K. The empirical distributions of the AD are displayed through the boxplots in Fig. 8. The distribution related to DZ method is the most shifted towards smaller values (with absolute deviations all smaller than 0.0045), followed by B, RD, and K, this latter being much more dispersed also in comparison to the normal case (Fig. 5).

Table 4 reports the MD and MAD values computed over the 441 scenarios for $k \in \{5, 6, 7, 8, 9, 10, 11\}$. We notice that the technique by Drezner and Zerom (2016) possesses by far the smallest MD in absolute value for any value k considered; the B estimates are characterized by a negative bias as, to a larger extent, the K estimates. The technique by Drezner and Zerom (2016) is the best performer also if we consider the MAD values, see Table 4; the method proposed by Kennan (2006) is the worst one, although the value of MAD quickly decreases when increasing k . The accuracy provided by the method proposed in Barbiero (2012) gets closer to that of Drezner and Zerom (2016) when k gets larger (see Fig. 9) as in the previous case.

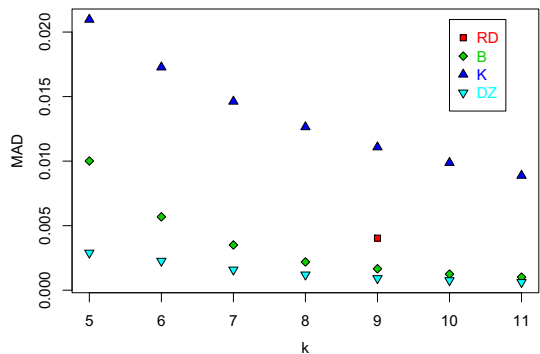
In Table 4, we reported also the values of MD and MAD for FORM. In terms of MAD, FORM is always overcome by the discretization methods (except for K) when $k \geq 6$; it is also characterized by large values of MD, which means that FORM, as occurs in the normal case, tends to overestimate the reliability parameter.

Along with the approximation accuracy, it is also important to assess and compare the computation time of the examined discretization methods, and of approximation by Monte Carlo simulation. Table 5 displays the computation times related to the 441 artificial scenarios investigated in the closeness study. One can notice how the method proposed by Drezner and

Table 4 Results of the comparative study for the hollow cylinder stress–strength problem with Weibull distributions; values of MD and MAD (to be multiplied by 10^{-3}) are displayed for FORM and for different discretization techniques and for different values of k

k	Method:	RD	B	K	DZ	FORM
5	MD		-2.566	-4.538	0.325	6.486
	MAD		10.794	20.862	3.075	7.200
6	MD		-1.616	-3.387	0.466	6.486
	MAD		5.221	17.422	2.158	7.200
7	MD		-0.496	-3.238	0.394	6.486
	MAD		3.157	14.663	1.451	7.200
8	MD		-0.477	-2.898	0.078	6.486
	MAD		2.110	12.676	1.065	7.200
9	MD	0.949	-0.636	-2.508	0.091	6.486
	MAD	4.332	1.785	11.209	0.950	7.200
10	MD		-0.485	-2.376	0.123	6.486
	MAD		1.280	9.887	0.744	7.200
11	MD		-0.371	-2.029	0.181	6.486
	MAD		1.063	8.845	0.626	7.200

Fig. 9 Values of MAD for the discretization techniques applied to the stress–strength problem with Weibull distributions. Each technique is depicted through a different symbol and colour. Note that the technique of Roy and Dasgupta (2002) for Weibull distribution is conceived for $k = 9$ only



Zerom (2016) is much more demanding than the other ones; this discrepancy becomes more severe moving from the Gaussian to the Weibull distribution.

4.3 Discussion

The closeness study presented in this section (other comparative studies have been carried out on other stress–strength problems, with similar results, which are not reported here for the sake of brevity) confirms that the discretization methods here investigated lead to different approximations and degree of accuracy when recovering the reliability parameter R , which is the final object of the investigation. Indeed, we can conclude that the method proposed in Drezner and Zerom (2016) performs better than the other methods in terms of both mean deviation and mean absolute deviation; this holds true both for the case of symmetrical (normal) and asymmetrical (Weibull) distributions. This method tries to accommodate two possibly conflicting requirements: (1) a k -point discretization of a continuous rv should be regarded as the process assigning to each of the k adjacent intervals, which its support can be split into, a representative value (namely, the mean of the continuous rv over that interval with respect

Table 5 Computation times in seconds for the FORM, the discretization techniques, and Monte Carlo simulation for completing the artificial scenarios considered in the closeness study

Distribution, method	k	FORM	GQ	RD	B	K	DZ	MC
Normal	5	1.0	1.8		0.9	1.0	159	1791
Weibull		3.3			2.3	2.4	818	2939
Normal	6	1.0	2.3		1.4	1.5	199	1791
Weibull		3.3			2.7	2.8	1274	2939
Normal	7	1.0	3.2		2.2	2.3	351	1791
Weibull		3.3			3.5	3.6	1799	2939
Normal	8	1.0	4.7		3.5	3.6	429	1791
Weibull		3.3			4.8	4.8	2383	2939
Normal	9	1.0	6.0	4.9	4.7	5.0	632	1791
Weibull		3.3		6.7	7.3	7.5	2906	2939
Normal	10	1.0	9.5		8.1	8.1	813	1791
Weibull		3.3			9.4	9.5	4059	2939
Normal	11	1.0	14.1		11.9	12.0	1046	1791
Weibull		3.3			13.5	13.4	4834	2939

to the original pdf) and its probability, (2) the approximating discrete distribution should be able to preserve the first two moments of the original continuous distribution. These requirements lead to a constrained non-linear optimization problem, which might adumbrate some computational hurdles, and actually, the technique becomes time-consuming as the number of discrete points increases: with $k > 9$ the computational effort becomes even larger than that required by MC simulation. Apart from this, no convergence problem is observed for deriving the final solution. On the contrary, some technical issues have been highlighted when applying the Gaussian quadrature method, which is the most acknowledged technique in the statistical literature. In this study, we observe that for asymmetrical distributions and moderately high number of approximating points it can easily fail in producing a feasible solution. The discretization proposal conceived by Kennan (2006) did not meet expectations: although the theoretical rationale behind the derivation of the approximating distribution is apparently clear and embraceable, and the applicability is universal, nevertheless the empirical results labelled it as the worst performer. However, one point that emerged from the comparative study is that it tends to improve its performance very quickly by increasing k , and this can make it a valid (even if not the best) alternative when one is interested in a straightforward approximation through a non-small number of points. The method in Barbiero (2012), which is a refinement of the methods of Roy and Dasgupta (2001, 2002), based on the matching of the cdf at the discrete points and some heuristic considerations, provides, similarly to Kennan (2006), a direct way to determine the points and probabilities of the approximating distribution without passing through the solution of an optimization or eigenvalue problem and with applicability to all continuous distributions. Its statistical performance has been proven to be satisfactory and close, especially for the highest values of k considered in the study, to that of Drezner and Zerom (2016).

The study also highlighted an important issue related to the choice of support points for discretized distributions. Some methods first choose the k support points and then assign the corresponding probabilities p_i in a second step: citing Woodruff and Dimitrov (2018) “The discretization process can be thought of as having two distinct components. The first

component is to select the points that will be used for discretization. Often, these points are chosen to be percentiles of the original distribution. The second component of discretization is to assign a probability mass to each point.” For example, in Roy and Dasgupta (2001) the k -point approximation of a standard normal is carried out by considering the integer points from $(k - 1)/2$ to $(k + 1)/2$: This choice may turn out to be not efficient, since, at least for high values of k , the continuous distribution is likely to be approximated through some extremely large/low and improbable values. In this sense, the Gaussian quadrature and the techniques proposed by Kennan (2006) and Drezner and Zerom (2016) seem to be more persuasive, since both points and probabilities are simultaneously (and not sequentially) elicited.

4.4 Software implementation

All the approximation and discretization techniques and the whole comparative study are implemented in the R programming environment through functions and codes that are available on request from the authors. The discretization method proposed by Drezner and Zerom (2016) is already available in the appendix of their paper. The Gaussian quadrature technique has been “translated” and adapted to R from the Matlab[®] code available in Toda (2020); in the original code, the matching of the first moments is implemented on a random sample drawn from a continuous distribution and not on the theoretical distribution. The FORM approximation is implemented through the homonym function `FORM` available in the R package `mistral` (Walter et al., 2016). In particular, we set the argument `Method` equal to `HLLRF`, standing for the standard Hasofer–Lindt–Rackwitz–Fiessler method (Hasofer and Lind, 1974; Rackwitz and Flessler, 1978) for recovering the design point iteratively, taking into account possible non-normality of random distributions, and the argument `IS` (Importance Sampling) equal to the default option `FALSE`.

Computation times for Monte Carlo simulation and the discretization and approximation techniques have been evaluated using RStudio Cloud, <https://rstudio.cloud/>.

The R functions implementing the techniques are conceived for the application to the normal and Weibull distributions involved in this study, but can be easily customized by the user with a small programming effort in order to enclose any other familiar parametric distribution. The data that support the findings of this study are available from the corresponding author upon reasonable request.

5 Final remarks

In several complex problems involving stochastic inputs, one or more continuous rv’s can be approximated by finite-support discrete rv’s in order to readily find a solution and/or make computations less time-consuming. There are many ways to accomplish this approximation; this paper reviewed several techniques employed in practice, trying to stress their peculiarities, and then apply them to an engineering problem within the statistical framework of stress–strength models. In the literature, a long-established and popular technique is moment equalization, which retains the values of the first moments of the original distribution and is however more manageable for symmetrical distributions and low values of the number of approximating points. A recently introduced proposal is intended to match only the first two moments, following a sort of latent variable approach for determining the discrete approximation. Other approaches preserve the value of the continuous cdf, or try to resemble it to some extent, at properly chosen discrete points, or find the best discrete approximation

by minimizing a distance between the continuous and the step-wise cdf, leading to a nice analytic solution which, by the way, provides uniform probabilities at the k points.

Through an empirical closeness study applied to a stress–strength model, involving either symmetrical (normal) or asymmetrical (Weibull) continuous distributions, these extant methods have been empirically compared in terms of their ability in correctly predicting the reliability parameter of the model. Moreover, we investigated the computation time required by the examined procedures, which is a relevant matter if one needs to apply them repeatedly, as it occurs when varying the experimental conditions to be simulated.

The discretization method proposed by Drezner and Zerom (2016), which is derived as the solution of a non-linear optimization problem ensuring the matching of the first two moments only of the original distribution, proves to be overall the best performer, in terms of mean (absolute) deviation; however, this method is much more time-consuming: differently from the other discretization methods, it embodies a constrained minimization, which has to be solved iteratively. Conversely, the method based on Kennan (2006), leading to equal-probability points, presents a poor global performance and more importantly exhibits a non-negligible bias when the value of the reliability parameter is very large (close to 1) or very small (close to 0). An approximate method specifically conceived for evaluating the reliability of a complex stress–strength model (the first-order reliability method, FORM) provides still acceptable solutions in a short computation time, but most discretization methods statistically outperform it in terms of accuracy, while preserving admissible computation times, especially when the number of discrete points gets larger.

We believe that this paper can be beneficial since differently from many scientific contributions that work out a statistical problem through a specific discrete approximation, here we are more concerned on how different approximation-by-discretization techniques and different values of the number of approximating points can affect the degree of dependability of the approximated solution itself. We focused on a single problem, but we provided also general recipes and hints for applying discretization to other contexts. Once more, we underline how the accuracy of an approximation has to be measured on the final outcome we are engaged in: in our work, the reliability parameter of a stress–strength model, whose independent continuous random components are (independently) discretized.

Future research will regard (i) other applications where discretization of continuous rv's is necessary or convenient, such as optimal portfolio problems (see an example in Toda, 2020) or approximate evaluation of aggregate risk, (ii) extension of discretization to correlated/dependent variables. This latter perspective can still concern the assessment of the reliability parameter of stress–strength models, if we assume that stress and strength (or their sub-factors) are no longer independent, a hypothesis which has already been explored in some recent works (see, e.g., Domma and Giordano, 2013) and which is contemplated also by FORM/SORM methods.

Acknowledgements The first author acknowledges the financial support to the present research by the Italian Ministry for Education, University and Research (FFABR 2017).

Funding Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory

regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barbiero, A. (2012). A general discretization procedure for reliability computation in stress–strength models. *Mathematics and Computers in Simulation*, 82, 1667–1676.
- Christofides, A., Tanyi, B., Christofides, S., Whobrey, D., & Christofides, N. (1999). The optimal discretization of probability density functions. *Computational Statistics & Data Analysis*, 31(4), 475–486.
- Clemen, R. C. (1991). *Making hard decisions: An introduction to decision analysis*. PWS-Kent Publishing Co.
- Colin, B., Dubeau, F., Khreibani, H., & de Tibeiro, J. (2013). Optimal quantization of the support of a continuous multivariate distribution based on mutual information. *Journal of Classification*, 30(3), 453–473.
- Cools, R. (2002). Advances in multidimensional integration. *Journal of Computational and Applied Mathematics*, 149(1), 1–12.
- Der Kiureghian, A. (2004). First-and second-order reliability methods. In E. Nikolaidis, D. M. Ghiocel, & S. Singhal (Eds.), *Engineering design reliability handbook*, chapter 14. CRC Press.
- D’Errico, J. R., & Zaino, N. A., Jr. (1988). Statistical tolerancing using a modification of Taguchi’s method. *Technometrics*, 30, 397–405.
- Domma, F., & Giordano, S. (2013). A copula-based approach to account for dependence in stress–strength models. *Statistical Papers*, 54(3), 807–826.
- Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and unsupervised discretization of continuous features. In *Machine learning proceedings* (pp. 194–202). Morgan Kaufmann.
- Drezner, Z., & Zerom, D. (2016). A simple and effective discretization of a continuous random variable. *Communications in Statistics-Simulation and Computation*, 45(10), 3798–3810.
- English, J. R., Sargent, T., & Landers, T. L. (1996). A discretizing approach for stress/strength analysis. *IEEE Transactions on Reliability*, 45, 84–89.
- Ghosh, T., & Roy, D. (2011). Statistical tolerancing through discretization of variables. *International Journal of Quality & Reliability Management*, 28(2), 220–232.
- Golub, G. H., & Welsch, J. H. (1969). Calculation of Gauss quadrature rules. *Mathematics of Computation*, 23(106), 221–230.
- Hammond, R. K., & Bickel, E. J. (2011). Discretization methods for continuous probability distributions. *Wiley encyclopedia of operations research and management science* (pp. 1–13).
- Hasofer, A. M., & Lind, N. C. (1974). Exact and invariant second-moment code format. *Journal of the Engineering Mechanics Division*, 100(1), 111–121.
- Kennan, J. (2006). A note on discrete approximations of continuous distributions. Department of Economics, University of Wisconsin.
- Kullback, S. (1959). *Information theory and statistics*. Wiley.
- Luceno, A. (1999). Discrete approximations to continuous univariate distributions—An alternative to simulation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(2), 345–352.
- Miller, A. C., & Rice, T. R. (1983). Discrete approximations of probability distributions. *Management Science*, 29(3), 352–362.
- R Core Team. (2019). A language and environment for statistical computing, version 3.6.0. Vienna, Austria. R Foundation for Statistical Computing.
- Rackwitz, R., & Flessler, B. (1978). Structural reliability under combined random load sequences. *Computers & Structures*, 9(5), 489–494.
- Roy, D. (2004). Discrete Rayleigh distribution. *IEEE Transactions on Reliability*, 53, 255–260.
- Roy, D., & Dasgupta, T. (2001). A discretizing approach for evaluating reliability of complex systems under stress–strength model. *IEEE Transactions on Reliability*, 50(2), 145–150.
- Roy, D., & Dasgupta, T. (2002). Evaluation of reliability of complex systems by means of a discretizing approach—Weibull set-up. *International Journal of Quality & Reliability Management*, 19(6), 792–801.
- Ruszczynski, A., & Shapiro, A. (2003). Stochastic programming. In A. Ruszczynski & A. Shapiro (Eds.), *Handbook in OR & MS* (Vol. 10, pp. 1–64). North-Holland Publishing Company.
- Šmid, M. (2009). The expected loss in the discretization of multistage stochastic programming problems—estimation and convergence rate. *Annals of Operations Research*, 165(1), 29–45.
- Smith, J. E. (1993). Moment methods for decision analysis. *Management Science*, 39(3), 340–358.
- Stroud, A. H. (1971). *Approximate calculation of multiple integrals*. Prentice Hall.

- Taguchi, G. (1978). Performance analysis design. *The International Journal of Production Research*, 16, 521–530.
- Tanaka, K., & Toda, A. A. (2013). Discrete approximations of continuous distributions by maximum entropy. *Economics Letters*, 118, 445–450.
- Tancrez, J. S., Chevalier, P., & Sema, P. (2011). Probability masses fitting in the analysis of manufacturing flow lines. *Annals of Operations Research*, 182, 163–191.
- Toda, A. A. (2020). Data-based automatic discretization of nonparametric distributions. *Computational Economics*, <https://doi.org/10.1007/s10614-020-10012-6>.
- Walter, C., Defaux, G., Iooss, B., & Moutoussamy, V. (2016). mistral: Methods in structural reliability analysis. R package version 2.1.0. <https://CRAN.R-project.org/package=mistral>.
- Woodruff, J., & Dimitrov, N. B. (2018). Optimal discretization for decision analysis. *Operations Research Perspectives*, 5, 288–305.
- Zhao, J., Han, C.-Z., Wei, B., & Han, D. (2012). A novel univariate marginal distribution algorithm based discretization algorithm. *Statistics & Probability Letters*, 82, 2001–2007.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.