Operations Research Perspectives 5 (2018) 288–305

Contents lists available at ScienceDirect

Operations Research Perspectives

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

Optimal discretization for decision analysis

Joshua Woodruff*, Nedialko B. Dimitrov

The University of Texas at Austin, Austin, TX, United States

ARTICLE INFO

Keywords:
Optimization
Discretization
Joint discretization

ABSTRACT

The use of discretization in decision analysis allows practitioners to use only a few assessments to estimate the certain equivalent (CE) or expected value of a decision without knowing the functional form of the distribution of each uncertainty. The discretization shortcuts are fast, but are created with a specific distribution, or families of distributions in mind. The discretizations are not formulated with the decision problem in mind. Each discretization is specific to one uncertainty distribution, or is even more generalized. In this article, we introduce a novel mathematical formulation for selecting an optimal discretization for a specific problem. With optimal discretization, a decision analyst can use the newly-created shortcuts in repeated decisions and improve the expected accuracy of the CE calculations.

1. Introduction

A common task in Decision Analysis is *discretization*, which involves reducing probability distributions of uncertainties to just a few point masses [13]. For many uncertainties in a Decision Analysis model, determining the value of multiple points in a distribution, let alone, the true distribution may be impossible or costly. Discretization is intended to reduce the cost of calculating certain equivalents (CEs) and simplify both communication and computation because it substitutes otherwise complex and computationally intensive integrations to the evaluations of just a handful of utilities. Discretizations significantly improve a decision analyst's ability to communicate with clients [16]. Even with an increase in computing power, discretizations allow for human-understandable assessment and evaluation of decisions.

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. An example of a discretization applied to two common distributions is shown in Fig. 1.

Because of the usefulness of discretization, many methods for discretization exist [13]. One discretizations method divides the probability distribution into intervals based on values or cumulative distributions. Each interval is given a percentile equal to it median or its mean. Finally, each interval is assigned a probability based on the size of the interval. Another way of discretizing is to choose percentiles and probabilities that match the moments of the original distribution. Typically, the underlying distributions of the uncertainties are unknown. As a result, these methods have given way to shortcut methods that are easy to implement and work across a broad range of distributions. Decision analysts seek discretization methods that produce near-correct certain equivalents across the entire class of decision problems of interest. Typically the number of percentiles that a decision analyst uses in a discretization is three, which provide a low, high, and most-likely value.

Though shortcut methods are easy to use, they are not created with the decision problem in mind. This can lead to reduced accuracy and situations where the results of different strategies are within a few percent of each other, sub-optimal decisions. Our method for choosing a discretization differs from these methods in that we seek a set of discretizations that provide the lowest certain equivalent error over a large set of potential distributions for a specific problem. The error particularly matters when the certain equivalent of a decision is close to zero. In a go, no-go decision, an error in the calculated certain equivalent would result in the wrong decision. We generate problem-specific discretizations by solving an optimization problem that chooses the percentiles and the corresponding probabilities that minimize an error metric over a large set of potential versions of a problem.

The main contributions of this article are: 1) We mathematically formulate the problem of selecting an optimal discretization for a set of decision problems. 2) We are able to derive tractable instances of this optimization problem. These tractable instances allow us to compute optimal discretizations for a specified set of decision problems. 3) Prior discretization methods produce discretizations for each uncertainty in

* Corresponding author.
E-mail addresses: joshua.woodruff@utexas.edu (J. Woodruff), ned.dimitrov@gmail.com (N.B. Dimitrov).
URLS: http://www.optimizedfinancialsystems.com/ (J. Woodruff), http://neddimitrov.org/ (N.B. Dimitrov).

https://doi.org/10.1016/j.orp.2018.09.002
Received 9 May 2017; Received in revised form 19 September 2018; Accepted 19 September 2018
Available online 20 September 2018

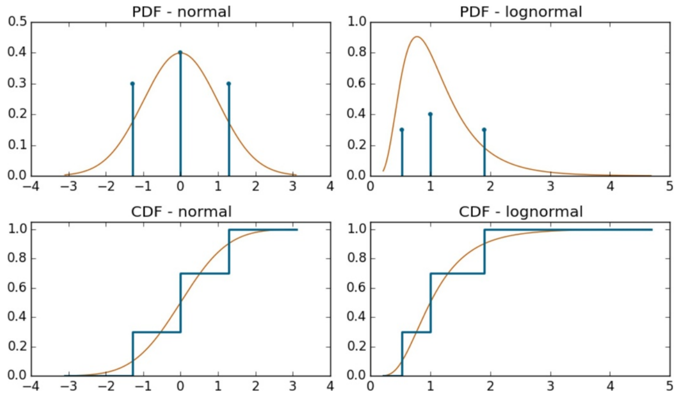

Fig. 1. These are two examples of the same three-point discretization applied to a standard normal and a log-normal distribution. The placement of the points on the independent axis are determined by the percentiles of the discretization, and the height is based on the probabilities assigned.

the decision problem. We introduce a *joint discretization* where the probabilities of the percentiles of a discretization are not independent from each other. The mathematical formulation for computing optimized discretizations opens a new area of computing joint discretizations. Finally, we show that our methodology for computing independent and joint discretizations outperforms prior methods across a set of computational examples.

This article is structured as follows. We begin with related work in Section 2, describing some popular and novel methods of discretization. Next we provide a general formulation for optimal discretization in Section 3. The general formulation is followed by the a tractable instance in Section 4.1 and modifications required to create a joint discretization in Section 4.2. In Section 5, we analyze two examples from the literature to determine the benefits of optimal and joint discretization, and the effects of sampling and pre-determining percentiles.We conclude with a discussion and future work in Section 6.

2. Related work

Discretizations are typically divided into distribution-specific methods and shortcuts. Distribution-specific methods require knowledge of the probability density function (PDF) of an uncertainty's distribution prior to discretization. These methods choose the discretization based on some criteria of the original distribution that the decision analyst is trying to match. Discretization shortcuts require experts to assess a few (usually three) percentiles of the uncertainty's distribution. They do not require knowledge of the shape or moments and are easily applied. A third type of method is a hybrid approach. This method assumes limited knowledge of the underlying distribution and provides a discretization based on this knowledge. Prior methods focus on computing discretizations independently for each uncertainty. In areas such as stochastic optimization there are more examples of multivariate discretizations [20]. We will review a few of these discretization methods and compare them to our approach.

In a decision analysis project, if the client, decision analyst, or some other expert knows the true distribution of each uncertainty, then the decision analyst can determine the CE of even the largest problem using Monte Carlo sampling or some other technique. With the true CE, the decision analyst can recommend the strategy with the highest CE with the certainty that this is the best recommendation. In reality, the form of each uncertainty is unknown. The decision analyst will elicit assessments for the uncertainties and apply a discretization to these uncertainties to create an estimate of the CE.

Two common distribution-specific discretization methods are bracket mean and Gaussian quadrature. Bracket mean, described by McNamee and Celona [15], is also known as the equal areas discretization. In bracket mean discretization, the PDF is partitioned into three sections, with probabilities 0.25, 0.50, and 0.25. The percentiles assigned to each region represent the mean value within the probability region. This method matches bounded means of the uncertainty's PDF. Though the method matches these bounded means for each uncertainty, there is no guarantee that the discretization with match the certain equivalent (CE) in the decision problem.

Miller III and Rice [16] and later Smith [19] proposed techniques based on Gaussian quadrature (GQ). With a discretization of N points, it is possible to match the first $2N-1$ moments of an uncertainty's distribution. The idea behind matching the moments comes from considering expectations of low degree polynomials. A discretization that matches an uncertainty in the first $2N-1$ moments, will produce the same expectation for any $2N-1$ order polynomial. Smith [19] improved the method for GQ by making it more efficient and provided examples of how GQ matched the moments of the input distributions ([19], Table 2, P.345). GQ requires knowledge of the distributions being discretized – at least its first $2N-1$ moments. If these are not known in the literature, one may require complex numerical integrations. An application of GQ also requires solving a multivariate system of polynomial equations, which is easily computed with matrix manipulation software. The polynomial matching argument misses cross-terms of the uncertainties when the utility depends on several uncertainties. From a client perspective, this discretization may ask for percentiles that are not easily assessed, such as the value of the uncertainty at the 99.50th percentile. In contrast, our method is able to limit the discretization's percentiles to easily assessed values, directly targets computing CEs, and the discretization is dependent on a set of decision problems, as opposed to a single specific distribution.

Shortcuts are easy to use and generally perform well. Shortcut discretizations do not require knowledge of the distributions of the uncertainties and apply the same discretization percentiles and probabilities to all uncertainties. Two common shortcut methods are the McNamee–Celona Shortcut (MCS) and the Extended Swanson–Megill (ESM) method. Both of these discretization methods use the 10th, 50th, and 90th percentiles. The Extended Swanson–Megill (ESM) method for discretization is described by Hurst et al. [10] and is commonly used in the oil and gas industries [2]. MCS assigns probabilities of (0.25,0.50,0.25) while ESM assigns probabilities of (0.3,0.4,0.3) to each of the respective percentiles. In examining several discretization methods, Keefer and Bodily [13] proposed the Extended Pearson–Tukey method. This method proposed using the 5th, 50th, and 95th percentile with probabilities of (0.185,0.630,0.185). Keefer and Bodily [13] found EPT outperformed several other methods. Additionally Hammond and Bickel [8] found EPT to have the best performance among MCS, EPT, and ESM when calculating average absolute error, average absolute percent error, maximum error, and maximum percent error. While easy to apply, these methods do not take into consideration specific knowledge about the decision problem. In contrast, our method computes an optimized discretization for a specific set of decision problems.

It is often known if the distributions of the uncertainties have specific shapes or are bounded. For example, when drilling for oil it can be assumed that the percentage of oil to be recovered will be between 0 and 100%. To that extent, there are newer shortcuts that are based on shape-specific assumptions. Hammond and Bickel [9] and Hammond and Bickel [8] offer new shortcuts based on the Pearson and Johnson families of distributions. Based on the specific zone of the assumed distribution of the uncertainty, Hammond and Bickel [8] provide a specific discretization to use. For example, a normal distribution is symmetric and unbounded on either side. A log-normal distribution is bounded from below, might be right-skewed, and unbounded from above. These methods are similar to the shortcut methods, but address a specific family of distributions. There are shortcuts for bounded, semi-bounded, and unbounded distributions. The purpose is to provide a better discretization that uses potentially available information while still allowing for an unknown distribution. Distribution-specific