Discretization, Simulation, and Swanson's (Inaccurate) Mean

J. Eric Bickel, SPE, and **Larry W. Lake**, SPE, University of Texas at Austin;
and **John Lehman,** Strategic Decisions Group

Summary

Swanson's mean (SM) is heavily used within the oil and gas industry to approximate continuous probability distributions such as the log-normal. In this paper, we document the errors induced by this practice, which, as we show, has no theoretical justification for any distribution other than the normal. In parallel, we review methods to discretize continuous distributions and compare these methods to Monte Carlo (MC) simulation. We demonstrate that the best discretization methods have an accuracy equivalent to that of tens of thousands of MC trials.

Introduction

Most decision or risk analyses include continuous random variables (e.g., oil in place, oil price, or porosity). Analysts are frequently concerned with how to best structure, compute, and communicate decision models under these circumstances. While decision trees are well suited for discrete random variables (RVs) with a few possibilities, they can become unmanageable when the number of outcomes is large. For example, trees may become too large to easily display or the large number of endpoints may require too many evaluations of a possibly costly (to evaluate) objective function (e.g., runs of a reservoir-simulation model). The question we address in this paper is how one should include continuous RVs in a decision tree given these limitations.

Two approximation methods, which are closely related, have been developed to address this problem: discretization and MC simulation. Specifically, suppose we are constructing a decision model that takes as an input the continuous RV X (e.g., oil in place). Under either discretization or MC simulation, we approximate the continuous probability-density function (PDF) $f(x)$, or the cumulative distribution function (CDF) $F(x)$, with a set of values $x_i \in X$, $i = 1, 2, \ldots, N$, and associated probabilities $p_i \equiv p(x_i)$. In MC simulation, N is generally large (in the thousands), and the x_i are drawn randomly from the entire support of X using methods such as the inverse cumulative method (Clemen and Reilly 2001). In this case, $X = F^{-1}(U)$, where F^{-1} is the inverse CDF and U is uniformly distributed over [0, 1]. The inverse cumulative method assures that X is distributed according to F. Discretization methods, on the other hand, comprise only a few points (three is common), but seek to choose points that will preserve desired properties of X. The most natural and common properties of interest are the raw (e.g., the mean) and central moments of X (e.g., variance, skewness, kurtosis). As we discuss in the following, if an approximation fails to preserve the moments of X, it is unlikely that it will accurately preserve the moments of the output distribution (Miller and Rice 1983; Smith 1993), including the mean, which are of interest to decision makers. Because the moments of a PDF do not uniquely determine the underlying PDF (an infinite number of distributions have a mean of zero and a variance of unity, for example), the discretization methods discussed in this paper, and used throughout the oil and gas industry, may fail to match particular percentiles of the PDF or extreme values even though they match many moments. This important and underappreciated distinction means that calculating the probability of exceeding particular values of the output variable (e.g., the probability of exceeding zero) on the basis of a cumulative distribution that comprises several discretized input uncertainties may not be justified. We do not address this issue here. Rather, our goal will be to match the moments of the input PDFs, with the caution that this process is designed only to estimate the moments of the output distribution.

Several discretization methods are in common use. For example, SM, which weights the 10th, 50th, and 90th percentiles of $F(x)$ by 0.30, 0.40, 0.30, is heavily used within the oil and gas industry (Megill 1984; Hurst et al. 2000; Rose 2001). However, SM is neither the only nor, as we shall see, the best choice.

Interest in discretization methods within the oil and gas industry has recently increased. For example, Arild et al. (2008) compared the use of SM to MC simulation in the context of a value-of-information problem. They demonstrated that the two methods yield different results, but they were unable to comment on which approach is closer to the true value of information because their problem did not have an analytic (closed-form) solution. Prange et al. (2009) compared a numerical approximation method to a discretization method and found that the discretization was inaccurate. Thus, these two papers have suggested the use of MC simulation instead of discretization. Willigers (2009), on the other hand, recommends the use of discretization instead of MC simulation to reduce the number of computations and speed the analysis of asset portfolios.

This paper is organized as follows. In the next section, we review the theory underlying discrete approximations and discuss the accuracy of common methods. In the third section, we discuss the accuracy of discretization methods and the demonstrably poor performance of SM. In the fourth section, we contrast the use of discrete approximations to MC simulation and demonstrate that the best discretization methods are equivalent to hundreds of thousands of MC trials. We provide a simple example in the fifth section that demonstrates that different discretizations yield different value estimates and possibly different recommendations for action. Finally, in the sixth section, we conclude and provide recommendations for practice.

Discrete Approximations: Methods and Shortcuts

Before we discuss specific discretization methods, we begin with a brief review and introduce the notation we use throughout the paper. The kth raw (or uncentered) moment of RV X with PDF $f(x)$ is

$$\mu_k = E[X^k] = \int_X x^k f(x)\mathrm{d}x \quad k = 0,1,2,\ldots, \quad \ldots\ldots\ldots\ldots\ldots (1)$$

where $E[\cdot]$ is the expectation operator. The zeroth raw moment specifies that the probabilities must integrate (or sum) to unity. The first raw moment μ_1 is the mean. Raw moments measure the distribution of an RV about the origin. Central moments, on the other hand, measure the distribution about the mean. (Central moments are simply raw moments of the transformed variable $Y = X - \mu_1$.) The kth central moment of X is

$$m_k = E[(X-\mu_1)^k] = \int_X (x-\mu_1)^k f(x)\mathrm{d}x \quad k = 0,1,2,\ldots. \quad \ldots (2)$$

The second central moment is the variance and is given the special symbol σ^2.

Original SPE manuscript received for review 5 October 2010. Revised manuscript received for review 3 February 2011. Paper (SPE 148542) peer approved 15 March 2011.

There is a one-to-one relationship between the raw and central moments. For example, if one knows the raw moments μ_k, the central moments can be found as follows (Papoulis 1984):

$$m_k = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} \mu_j \mu_1^{k-j}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Hence, the second central moment, the variance, expressed in terms of the raw moments, is $m_2 = \mu_2 - \mu_1^2 \equiv E[X^2] - (E[X])^2$.

Two normalized central moments, skewness and kurtosis, measure other properties of the distribution. Skewness is a measure of asymmetry and is given by $\gamma_3 = m_3 m_2^{-3/2}$. Kurtosis measures the distribution's degree of "peakedness" or how thick its tails are. Kurtosis is defined as $\gamma_4 = m_4 m_2^{-2} - 3$. The minus three normalizes kurtosis relative to the normal distribution, which has a kurtosis of three.

The Motivation for Moments. Suppose we are interested in a value function $v(x)$, which is a function of an RV X. This value function might be net present value (NPV) or ultimate hydrocarbon recovery. If this value function is sufficiently differentiable, such that it can be approximated by a polynomial $P(x)$ of degree N, then we have

$$v(x) \approx P(x) = \sum_{k=0}^{N} w_k x^k, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

for the weights w_k. [This discussion closely follows Smith (1993).] The mean, or expected value of $v(x)$ is given by

$$E[v(X)] \approx E(P(X)) = E\left[\sum_{k=0}^{N} w_k X^k\right] = \sum_{k=0}^{N} w_k E(X^k). \quad \ldots\ldots (5)$$

Thus, the expected value of $v(x)$ is an expansion of the raw moments of X. Therefore, if one wants to accurately compute the expected value of the value function (e.g., expected NPV), *all* the moments of $f(x)$ must be represented accurately. It is not sufficient, as is sometimes assumed, to accurately reflect only the mean value of input uncertainties. To make this concrete, suppose that $v(x) = x^2$ and that X is a normal RV with zero mean and unit variance. The expected value of $v(x)$ is equal to unity, while v evaluated at the expectation is equal to zero. Our goal then will be to develop discretization methods that preserve the moments of input distributions. Before we discuss particular discretizations, however, we will formally state the approximation problem.

Formal Statement of Approximation Problem. Given a PDF $f(x)$ and function $\Omega(x)$, we wish to approximate a definite integral with a finite sum,

$$\int_X f(x)\Omega(x)\mathrm{d}x \approx \sum_{i=1}^{N} p_i \Omega(x_i), \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

by choosing values x_i and weights (or probabilities) p_i.

In other words, we seek to approximate the PDF with a discrete probability mass function (PMF). This procedure, originally developed by Gauss in the nineteenth century (Gauss 1866), is known as Gaussian quadrature. If Ω is a polynomial then the approximation in Eq. 6 holds with equality (Davis and Rabinowitz 1984).

In this paper, we analyze the particular case $\Omega(x_i) = x_i^k$, where $k = 0, 1, 2, \ldots$ represents the kth raw moment of X. Thus, we have

$$E[X^k] = \int_X f(x)\Omega(x)\mathrm{d}x = \int_X f(x)x^k\mathrm{d}x = \sum_{i=1}^{N} p_i x_i^k \quad \text{for} \quad k = 0,1,2,\ldots,N, \quad \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

for a set of probabilities p_i and values x_i; our challenge is to find these parameters. As an example, consider $k = 0$, in which case Eq. 7 requires that the discrete probabilities sum to unity. When $k = 1$, Eq. 7 requires that we match $E[X]$, or the mean, and so on. This application of Gaussian quadrature is referred to as *moment matching* (Miller and Rice 1983; Smith 1993). An approximation with N points can match $2N$ moments of X, including the zeroth moment (Stroud and Secrest 1966; Miller and Rice 1983). For example, a three-point approximation can match the zeroth through the fifth moment of X. This remarkable fact enables one to preserve many characteristics (mean, variance, skewness, kurtosis) of commonly used distributions with a limited number of points, assuming these moments are finite.

Gaussian quadrature provides both an organizing framework for understanding the objective of discretization and a computational method for determining the best approximation. However, many commonly used approximations fail to match some or all of the moments of X. We discuss the differing approaches in the following subsections, which we divide into general methods and shortcuts. The general methods are intended for use on a case-by-case basis and are tailored to the underlying PDF. The shortcuts are nonparametric in that they are discretizations that are applied to any PDF, irrespective of its shape. Typically, these are discretizations that were found to work reasonably well over some set of commonly encountered PDFs but are applied more generally than might be warranted.

General Discretization Methods. In this subsection, we describe three general discretization methods: moment matching, bracket median, and bracket mean.

Moment Matching. The most accurate general method for matching moments is, of course, Gaussian quadrature or moment matching itself. (While we are using a special case of Gaussian quadrature, where $\Omega(x_i) = x_i^k$, we will use the terms moment matching and Gaussian quadrature interchangeably.) This approach can be applied to any PDF for which the moments are known. Quadrature rules have been calculated and published for several commonly encountered PDF families. For example, Miller and Rice (1983) presented the two-, three-, and four-point Gaussian quadratures for the standard uniform, normal, and exponential distributions. We repeat these results in **Table 1.** [Miller and Rice provided their results to six significant figures, but we round these to three significant figures to facilitate communication. Readers requiring greater accuracy should consult Miller and Rice (1983).] However, instead of providing the values x_i we provide the percentiles of the excess distribution function (EDF), also called the complementary CDF, which is used more frequently than the CDF in oil and gas settings. These percentiles are $\alpha_i \equiv G(x_i) \times 100 = [1 - F(x_i)] \times 100$.

We summarize an approximation as the set of probabilities and percentiles (p_i; α_i). As shorthand, we will refer to the Zth percentile of the EDF as the PZ. For example, the 50th percentile is the P50. The two-point approximation of the normal distribution may be summarized as (0.500, 0.500; P84.1, P15.9). The three-point approximation of the normal is (0.167, 0.667, 0.167; P95.8, P50.0, P4.2). This simple approximation will match the first six moments of the normal distribution, including the requirement that the probabilities sum to unity. It may help the reader to think of the discretizations in Table 1 as probability trees. For example, the three-point approximation for the normal is shown in **Fig. 1.**

In the case of the uniform distribution, Gaussian quadrature does not result in equal weights on uniformly dispersed values. For example, one might think that an equal weighting of the P75, P50, P25 would perfectly match the uniform distribution because both the values and the probabilities have been divided uniformly. In fact, this approximation underestimates the variance by 50%.

Bracket Median and Bracket Mean (or Equal Areas). In two intuitively appealing methods, bracket median (Clemen and Reilly 2001) and bracket mean (McNamee and Celona 1990), the excess distribution $G(x)$ is divided horizontally into N intervals, as shown in **Fig. 2.**

It is common for the intervals to be of unequal size in the bracket-mean approach and of equal size in the bracket-median approach, although in neither case is it required. The probability that X will be in Interval i is $G_i - G_{i-1}$. Given that X is in Interval i, we summarize the conditional distribution of X (conditional on X being within the interval) by a single number. Bracket median takes

TABLE 1—TWO-, THREE-, AND FOUR-POINT GAUSSIAN QUADRATURE FORMULAE FOR COMMON DISTRIBUTIONS

Distribution	Two Points		Three Points		Four Points	
	p_i	α_i	p_i	α_i	p_i	α_i
Uniform	0.500	78.9	0.278	88.7	0.174	93.1
$f(x) = 1$	0.500	21.1	0.444	50.0	0.326	67.0
$0 \le x \le 1$			0.278	11.3	0.326	33.0
					0.174	6.9
Normal	0.500	84.1	0.167	95.8	0.046	99.0
$f(x) = (2\pi)^{1/2} Exp[-\frac{1}{2}x^2]$	0.500	15.9	0.667	50.0	0.454	77.1
$-\infty \le x \le \infty$			0.167	4.2	0.454	22.9
					0.046	1.0
Exponential	0.841	55.7	0.711	66.0	0.603	96.8
$f(x) = Exp[-x]$	0.146	3.3	0.279	10.1	0.357	17.5
$x \ge 0$			0.010	0.2	0.039	1.1
					0.001	0.01

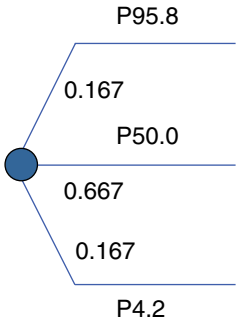

Fig. 1—Three-point Gaussian quadrature for the normal distribution.

this number to be the median, while bracket mean uses the mean. Because these conditional distributions are generally skewed, the median and the mean differ and the two approaches may result in (very) different discretizations. For example, applying the three-point bracket-median and bracket-mean methods with intervals of 0.25, 0.50, 0.25 to the normal distribution yields discretizations of (0.25, 0.50, 0.25; P87.5, P50.0, P12.5) and (0.25, 0.50, 0.25; P89.8, P50.0, P10.2), respectively. The difference between the approaches is greater for skewed distributions such as the log-normal.

The advantage of the bracket-median approach is that the median of each interval is simply that interval's midpoint, $G_{i-1} + (G_i - G_{i-1})/2$, which can be read directly off of the excess distribution, whereas, the mean of each interval must be calculated. For certain distributions, such as the normal, this calculation is straightforward. For other distributions, including subjectively assessed distributions that may not belong to any family, this calculation may be complex. However, there is a graphical method to estimate the interval means that gives bracket mean its more common name of equal areas. As shown in Fig. 2, the mean of each Interval i is the point at which the area A_i to the left and above the excess distribution is equal to the area below and to the right. The interested reader should see McNamee and Celona (1991, pp. 18–21) for a proof of this fact. In the first author's experience, individuals are skilled at finding these points by eye.

Discretization Shortcuts. The general methods discussed previously can be tailored to the shape of the distribution. For example, the intervals in equal areas need not be symmetric and extra attention can be placed on particular portions of the distribution, such as the tails. Yet, this flexibility comes at a cost—one must compute the

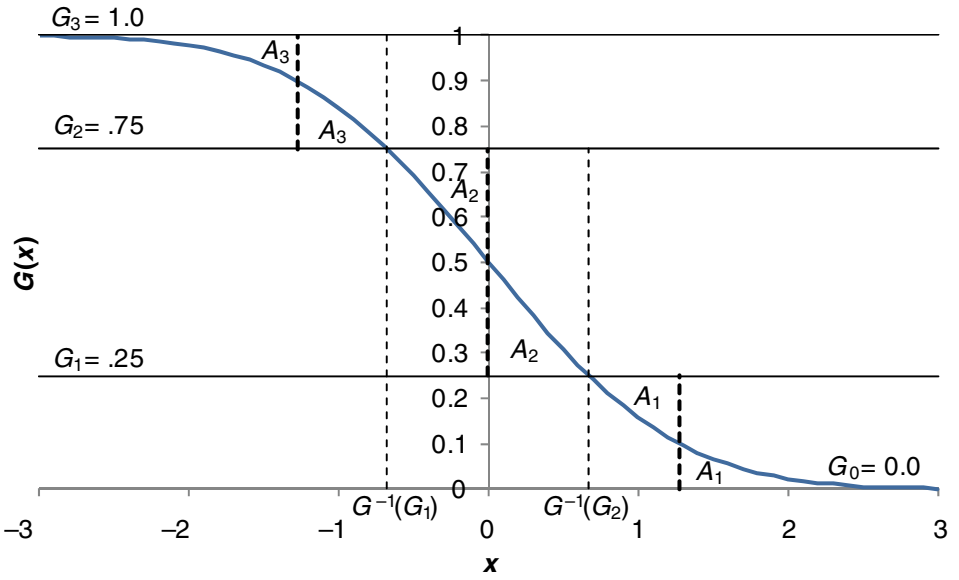

Fig. 2—Illustration of equal-areas discretization method.

TABLE 2—THREE-POINT DISCRETIZATION SHORTCUTS											
EPT		MCS		ESM		MRO		GQN		ZDT	
p_i	α_i	p_i	α_i	p_i	α_i	p_i	α_i	p_i	α_i	p_i	α_i
0.185	95.0	0.250	90.0	0.300	90.0	0.248	91.5	0.167	95.8	0.333	89.0
0.630	50.0	0.5000	50.0	0.400	50.0	0.504	50.0	0.667	50.0	0.333	50.0
0.185	5.0	0.250	10.0	0.300	10.0	0.248	8.5	0.167	4.2	0.333	11.0

appropriate discretization. For this reason, several shortcut methods have been developed. In what follows, we discuss the approximations in the order in which they were originally proposed.

Pearson and Tukey (1965) developed a three-point discretization meant to closely approximate the mean of common PDFs, including the normal, beta, gamma, inverse gamma, and Student's *t* distributions (the uniform was not included). Their three-point approximation of the mean is (0.185, 0.630, 0.185; P95, P50, P5), which was found to work well except for distributions that were highly skewed. Pearson and Tukey did not recommend this approximation for higher moments and instead provided a more complex approximation for the standard deviation. However, Keefer and Bodily (1983), on the basis of work by Keefer and Pollock (1980) and their own supporting analysis, suggested treating the Pearson-Tukey approximation as a complete PMF and referred to this as extended Pearson-Tukey (EPT).

The equal-areas method discussed previously was developed by Jim Matheson and his colleagues at the Stanford Research Institute (SRI) between the late 1960s and the early 1970s (personal communication with Jim Matheson and Peter McNamee). Application of this method to the normal distribution produces an approximation of (0.25, 0.50, 0.25; P89.8, P50, P10.2). On the basis of this, SRI began using a shortcut of weighting the P90, P50, P10 by 0.25, 0.50, 0.25, which is sometimes referred to at the 25-50-25 approximation. This method was then heavily used and popularized by Strategic Decisions Group (SDG), which was founded by individuals from SRI's decision analysis group. SDG trained hundreds of oil and gas professionals in decision analysis methods and helped to establish existing decision analysis programs at several major corporations, including Chevron; this explains the use of 25-50-25 in oil and gas settings. The shortcut is described in McNamee and Celona (1990, pp. 32–33)) and has come to be known as the McNamee and Celona shortcut (MCS). McNamee and Celona, SDG consultants at the time, cautioned that this shortcut should be used only in the early stages of analyzing a decision and that one needs to carefully assess the distribution and develop a full discretization (using equal areas) "more carefully later on! [emphasis in original]" (McNamee and Celona 1990). Over time, this guidance has been widely forgotten, and today MCS is commonly applied without regard for the shape of the underlying distribution and is not followed with a secondary and more careful assessment and discretization.

While working for Exxon, Roy Swanson, in a 1972 internal memo, proposed approximating mean reserves by weighting the P90, P50, P10 of the reserves EDF by 0.30, 0.40, 0.30 (Megill 1984 in Appendix B; Hurst et al. 2000). Following the preceding discussion, we will refer to this as a 30-40-30 weighting. According to Megill (1984), Swanson arrived at this rule empirically and found that it reasonably approximated the mean of modestly skewed distributions. Like Pearson and Tukey before him, Swanson (apparently) did not propose using his approximation to estimate higher moments or as a complete PMF. However, Keefer and Bodily (1983) proposed treating Swanson's 30-40-30 rule as a complete PMF and referred to it as extended Swanson-Megill (ESM). The general use of ESM, especially in the case of the log-normal, has been advocated by Pete Rose and his colleagues at Rose & Associates (Hurst et al. 2000; Rose 2001). As we will see later, ESM is close to a Gaussian quadrature for the normal distribution that matches the mean and variance. However, when directly applied to a log-normal distribution, ESM fails to match the mean and *significantly* underestimates the variance and the skewness.

Before discussing the other discretization methods, we pause briefly so that we can emphasize the difference between developing an approximation of the mean and an approximation of the PDF. A mean does not uniquely determine a PDF; there are an infinite number of PDFs with the mean of 5, for example. Therefore, one could very easily find an approximation that matches the mean, but fails to faithfully represent the underlying PDF. This is the case with SM. Swanson did not suggest using his method to approximate the PDF of reserves; he only suggested that it be used to approximate the mean. However, approximations that work well only for the mean are not particularly useful in most decision analyses. Using a discretization in a probability or decision tree as shown in Fig. 1 implicitly assumes that it is an accurate representation of the PDF, not just the mean of that PDF. Thus, strictly speaking, Swanson's method cannot be used in a decision tree because it was intended merely to be a method for estimating the mean of a distribution. When Swanson's values (0.30, 0.40, 0.30; P90, P50, P10) are used in a decision tree, it is really an application of ESM that describes a PMF.

Returning to our discussion of the various shortcuts, Miller and Rice (1983) introduced the use of Gaussian quadrature to the decision analysis literature. As detailed earlier, this method can exactly match as many moments of any PDF as desired, as long as the moments are finite. However, in practice it would be helpful to have these discretizations precalculated. This is possible in the case of known PDF families, as shown in Table 1. When one is dealing with a distribution that is not from a known family (as might happen if the distribution is directly assessed by an expert, for example), Miller and Rice proposed several generic discretizations based on Gaussian quadrature. For a three-point approximation, Miller and Rice proposed weighting the P91.5, P50.0, and P8.5 by 0.248, 0.504, 0.248. This approximation has become known as the Miller-Rice one step (MRO). Notice that MRO is very close to the bracket-mean approximation applied to a normal distribution and MCS. This correspondence further supported MCS and SDG's use of it.

D'Errico and Zaino (1988) and Zaino and D'Errico (1989) used Taguchi's method (Taguchi 1978) to develop two approximations. The first equally weighs the P89, P50, P11, which we refer to as the Zaino-D'Errico-Taguchi (ZDT) approximation. The second applies the three-point Gaussian quadrature formula for the normal distribution (0.167, 0.667, 0.167; P95.8, P50.0, P4.2), displayed in Table 1, more generally. We will refer to this approximation as GQN. We summarize each of the discretization shortcuts in **Table 2.**

Moment Matching With Fixed Values or Fixed Probabilities. After many years of use, approximations with values fixed at the P90, P50, P10 (MCS and ESM) or weightings of 25-50-25 (MCS) or 30-40-30 (ESM) have become common. Yet, the ESM and MCS shortcuts are not distribution specific and, therefore, may induce unnecessary errors. In this subsection, in an effort to improve practice, we apply moment matching to develop rules for (a) weighting the P90, P50, and P10 and (b) values with fixed weights of 25-50-25 or 30-40-30. We consider the uniform, normal, exponential, and triangular distributions. These results appear in **Tables 3 through 5.**

As seen in Table 3, the weights for the normal are nearly identical to ESM. As such, ESM almost matches the mean and variance of a normal distribution. The listed triangular results are for a mode, *c,* of 0.5. However, the weights are a very weak function of *c*. For example, when $c = 0.1$ the weights are (0.277, 0.451, 0.272). Thus, the values listed in Table 3 should provide satisfactory results in most situations.

It is not always possible to fix the values and find a feasible solution (Smith 1993). For example, a weighting of a log-normal's

TABLE 3—THREE-POINT MOMENT-MATCHING WEIGHTS FOR P90, P50, P10 VALUES

Uniform	Normal	Exponential	Triangular*
$f(x)=1$	$f(x)=(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}x^2}$	$f(x)=e^{-x}$	$f(x)=\begin{cases}\frac{2(x-a)}{(b-a)(c-a)} & a\le x\le c\\ \frac{2(b-x)}{(b-a)(b-c)} & c<x\le b\\ 0 & \text{otherwise}\end{cases}$
$x\in[0,1]$	$x\in(-\infty,\infty)$	$x\ge 0$	$x\in[a=0,b=1,c=0.5]$
0.260	0.304	0.465	0.273
0.480	0.392	0.175	0.454
0.260	0.304	0.360	0.273

*Note: Triangular discretization is a function of *c*, but this relationship is weak enough to ignore.

TABLE 4—THREE-POINT MOMENT-MATCHING VALUES FOR 25-50-25 WEIGHTS (SAME PDFs AS TABLE 3)

Uniform	Normal	Exponential	Triangular
P75.0	P92.1	P93.3	P78.9
P62.1	P50.0	P44.6	P50.0
P0.9	P7.9	P11.2	P21.1

TABLE 5—THREE-POINT MOMENT-MATCHING VALUES FOR 30-40-30 WEIGHTS (SAME PDFs AS TABLE 3)

Uniform	Normal	Exponential	Triangular
P74.9	P90.2	P91.8	P76.4
P64.0	P50.0	P42.3	P50.0
P6.5	P9.8	P9.5	P23.6

P90, P50, P10 that will simultaneously match the mean and variance is very often impossible to find if the underlying distribution is even modestly skewed. This underscores the futility of the common practice of applying ESM or MCS without regard for the underlying distribution or, indeed, instead of considering the nature of that distribution. To remedy this, in **Table 6** we present several approximations for the log-normal with any mean but the listed standard deviation (in terms of ln *X*), using four points. These approximations will match the mean and variance when applied directly to log-normal distributions with the stated standard deviation; as we discuss more fully later, matching the skewness of a log-normal is very difficult. The values that we have found to work well are the P90, P50, P5, and the mode (most likely value). If *X* is log-normally distributed, then the mode of *X* is $EXP[\mu-\sigma]$, where $\mu = E[\ln X]$ and $\sigma^2 = E[((\ln X)-\mu)^2]$.

Accuracy of Discretization Methods

Given the variety of discretization methods, the question naturally arises as to which approximation is best. Or, perhaps more correctly, in which situations do the various approximations perform well or poorly? When the PDF is from a family given in Table 1 (or Tables 3–6), the Gaussian quadrature formulas exactly match the first 2*N* underlying moments of the original PDF, by definition. Thus, they can be taken as the highest standard of accuracy.

Miller and Rice (1983) proved that equal areas will always underestimate the even moments of the original distribution. This occurs because x^k is convex when *k* is even and, by Jensen's inequality ($E[v(X)] \ge v(E[X])$, if *v* is a convex function.), the expectation of a convex function will exceed the value of the function evaluated at the expectation. In the case of odd *k,* the direction of the error is more difficult to sign. If *x* is positive (negative) then the odd moments will be underestimated (overestimated), since x^k is convex (concave). Thus, direct application of equal areas to a log-normal distribution (which cannot take negative values) will underestimate all moments.

Keefer and Bodily (1983) tested EPT and ESM (among others) across a range of beta distributions, which can assume a wide variety of shapes and thus higher moments. They concluded that EPT is the "clear winner." Both EPT and ESM approximated the mean reasonably well; the average (maximum) errors were 0.02% (0.07%) and 0.05% (0.33%) for EPT and ESM, respectively. However, the two approaches differed in their ability to approximate the variance. The average (maximum) errors were 0.46% (–1.6%) and 2.7% (11.1%). Keefer (1994) extended the analysis of Keefer and Bodily (1983) by analyzing the accuracy of MCS, MRO, GQN, and ZDT across a range of beta distributions. In terms of estimating the mean and variance, he found that EPT slightly outperformed GQN, and that they both dominated ESM, MRO, MCS, and ZDT.

Swanson's (Inaccurate) Mean. Given the widespread use of ESM in the oil and gas industry, and the fact that it is being used as intended, which is not the case for MCS, it seems appropriate to scrutinize its accuracy. (The reader should bear in mind that MCS was never recommended as a final approximation, as ESM has been.) The studies discussed in the preceding demonstrated the rule's inability to accurately approximate the variance of many

TABLE 6—REPRESENTATIVE FOUR-POINT DISCRETIZATIONS FOR LOG-NORMAL

σ = .25		σ = .50		σ = .75		σ = 1.0	
γ_3 = 0.78		γ_3 = 1.75		γ_3 = 3.26		γ_3 = 6.18	
p_i	α_i	p_i	α_i	p_i	α_i	p_i	α_i
0.203	90.0	0.151	90.0	0.186	90.0	0.436	90.0
0.288	59.9	0.265	69.1	0.235	77.3	0.274	84.1
0.303	50.0	0.378	50.0	0.357	50.0	0.017	50.0
0.206	5.0	0.206	5.0	0.222	5.0	0.272	5.0

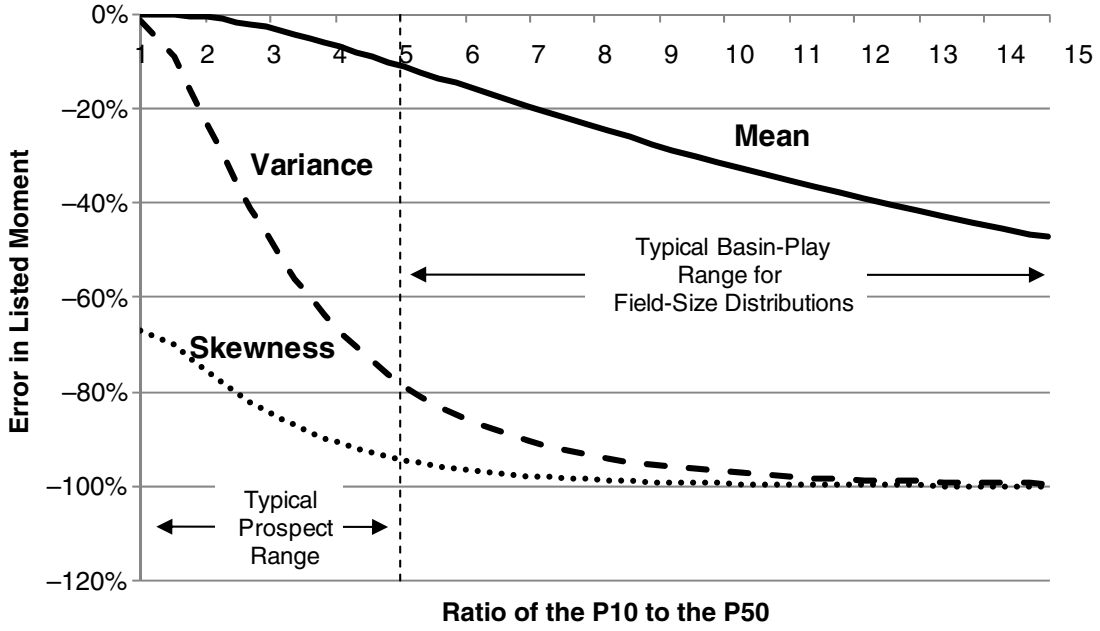

Fig. 3—Error in SM with direct application to log-normal distribution. Compare with Megill (1984).

distributions. In this section, we more carefully analyze the use of ESM to approximate log-normal distributions; since its inception, ESM has been used to summarize reserves distributions, and reserves are widely held to be log-normally distributed (Rose 2001). Likewise, Rose has advocated the use of ESM specifically for use with log-normal distributions (Rose 2001 in Appendix B).

Megill (1984) directly applies ESM to the log-normal and finds that ESM underestimates the mean by approximately 10% for modestly skewed distributions, which he associates with "typical prospect ranges." However, he also finds that ESM underestimates the mean by 45% for more-skewed distributions, which he associates with "typical basin-play ranges for field size distributions." Megill concludes that "Swanson's rule should not be applied to obtain the mean of play or basin assessments." We extend Megill's analysis in **Fig. 3,** which plots the error in ESM's estimate of the mean, variance, and skewness for a log-normal distribution, against the ratio of the P10 to the P50, which Megill intended to be a measure of skewness. As Megill stated, ESM underestimates the mean by up to 45% in this example. What Megill did not mention is that it also underestimates the variance by 80% for the typical prospect range and by 100% for more-skewed distributions. ESM's estimation of the skewness is even worse.

Rose (2001) supports his use of ESM by arguing that reserves above the P1 of a log-normal occur with much less than a 1% frequency, and therefore the log-normal should be truncated above this point. Doing so reduces skewness and does improve the accuracy of ESM. However, Rose supports his argument by analyzing only the mean and examining a single log-normal distribution with a mean of 15.1 and a standard deviation of 28.2—implying a skewness of 3.9. In this case, ESM underestimates the true mean by 1.5%. However, it also underestimates the variance by 59% and the skewness by 78%. Furthermore, under a different truncated log-normal distribution with a skewness of 4.9, ESM underestimates the mean by 10%, the variance by 77%, and the skewness by 83%.

SM can be directly applied to log-normal distributions by applying it to the *logarithm* of X, instead of directly to X. If X is log-normally distributed, then ln X is normally distributed and the moments of X are functions only of the mean μ and variance σ^2 of ln X. The equations for the first four moments of X appear below:

$$\mu_1(X) = \exp\left[\mu + \tfrac{1}{2}\sigma^2\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

$$m_2(X) = \exp\left[2\mu + \sigma^2\right]\left(\exp\left[\sigma^2\right] - 1\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots (9)$$

$$\gamma_3(X) = \left(\exp[\sigma^2] + 2\right)\left(\exp[\sigma^2] - 1\right)^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (10)$$

$$\gamma_4(X) = \exp\left[4\sigma^2\right] + 2\exp\left[3\sigma^2\right] + 3\exp\left[2\sigma^2\right] - 6. \quad \ldots\ldots (11)$$

To use SM, we would simply determine μ and σ^2 by applying the 30-40-30 approximation to ln X, which is normal. Substituting these values into Eqs. 8 through 11 would almost perfectly match the first four moments of X, which is log-normal. This approach would reduce all the errors in Fig. 3 to zero. This log-SM approximation is a significant improvement over SM and is only slightly more complicated—requiring the use of a logarithm and of an exponential.

If we want to truncate the log-normal distribution at, say, the P1, we simply need to truncate the underlying normal distribution at the P1, which is elementary. Then we can find a Gaussian quadrature for this truncated normal. For example, if we want to truncate the log-normal at the P1, a Gaussian quadrature of (0.32, 0.37, 0.31; P89.1, P50, P9.9) of the original untruncated normal will exactly match the first three moments of the normal and, therefore, the log-normal distribution. This approximation would reduce all of the errors in Rose's (2001) example to zero.

Why Not Just Simulate?

Thus far, we have compared discretization methods among themselves and found that some, such as ESM and naïve uses of MCS, produce significant errors. One may wonder then, given the widespread availability and use of MC methods, why use discretization at all? (As one analyst argued, "After all, my Monte Carlo software package includes the log-normal distribution.") We must remember that MC is also an approximation. Discretization methods induce *approximation error*, while MC methods include *sampling error*. The relevant question is whether or not MC is more accurate than discretization and which situations lend themselves to each method. To address this, this section offers some reasons as to why one may prefer to use discretization rather than simulation. More importantly, we determine how many MC samples are required to achieve the same accuracy as each discretization method.

Modeling Difficulties. From a modeling perspective, MC simulation is not easy to implement in situations that include downstream decisions or options. A generic example appears in **Fig. 4.** In this case, a decision $D1$ is made at the beginning of Stage 1. Then,

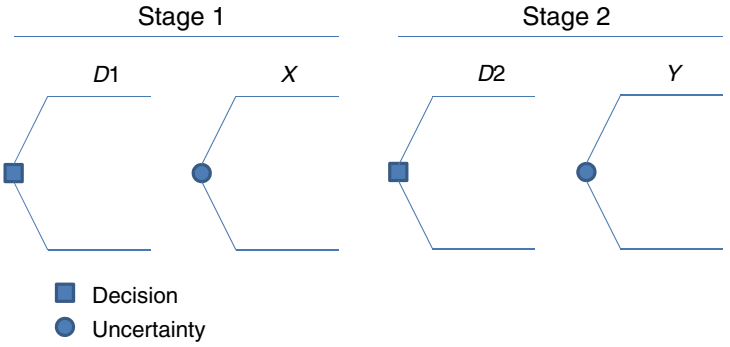

Fig. 4—Example of decision situation that may not lend itself well to MC methods.

a possibly continuous uncertainty X is revealed, yielding a particular value x. At the start of Stage 2, decision $D2$ is made, with knowledge of $D1$ and x. Then, the possibly continuous uncertainty Y is realized. To make this situation concrete, suppose that $D1$ is a decision about the acquisition of seismic data, X is the result of the seismic survey, $D2$ is the decision to drill, and Y is the economic value of the reservoir (Stibolt and Lehman 1993; Bickel et al. 2008).

To evaluate this decision tree, we start at the end and roll back (McNamee and Celona 1990; Clemen and Reilly 2001). Suppose we only seek to maximize expected values (i.e., we are risk neutral) and that X and Y are dependent. In this case, we would need to compute $E[Y \mid X = x]$ to choose the optimal alternative at Stage 2. If this conditional mean cannot be expressed as a function of x, we must simulate Y for each realization of X. If we are performing 1,000 MC trials, for each of the 1,000 trials for X, we would have to perform an additional 1,000 trials for $Y \mid X = x$, requiring $1{,}000^2$ or 1 million trials. This nested MC simulation is not straightforward to implement and may require a long time to evaluate.

Costly Evaluation. The second reason one may prefer discretization methods is that they require fewer points (generally *many* fewer points) to match the underlying moments. Because each point or MC trial requires the evaluation of an output function (e.g., NPV), this process may be computationally costly. To understand this more fully, we can determine *the number of MC samples that would be required to achieve the same accuracy as a discretization method* (Pfeifer et al. 1991). We refer to this as S-equivalence.

The N-point Gaussian quadratures in Table 1 exactly match the first $2N$ moments of the uniform, normal, and exponential distributions. Even in the case when $N = 2$, these approximations exactly match the mean, variance, and skewness of the original distributions. As such, MC could not do a better job approximating these moments. The shortcuts listed in Table 2 may, on the other hand, fail to match some moments of certain distributions. We will determine S-equivalents for these shortcut-distribution combinations. On the basis of Eq. 5 and for ease of exposition, we focus only on the raw moments.

Uncertainty in Moments. Imagine creating a new distribution $Y_k = X^k$ from which we will sample in order to estimate the raw moments of X. The mean of Y_k, μ_k, is the kth raw moment of X. For example, if $k = 1$, then we would simply sample from X. Each set of S samples would produce an estimate of the mean of X. However, each time we rerun our MC simulation using S samples, we would compute a different mean. In other words, the sample raw moments are random variables. We quantify their uncertainty by computing their central moments (e.g., the variance). The jth central moment of Y_k is

$$m_j^k \equiv E[(Y_k - E[Y_k])^j] = E[(X^k - \mu_k)^j]. \quad \ldots\ldots\ldots\ldots\ldots\ldots (12)$$

If k equals 1 and j equals 2, then we have

$$m_2^1 \equiv E[(X - \mu_1)^2] = \mu_2 - \mu_1^2, \quad \ldots\ldots\ldots\ldots\ldots\ldots (13)$$

which is the familiar formula for the variance written in terms of raw moments. When referring specifically to the variance of the kth raw moment, we will use the expression $\sigma_k^2 \equiv \sigma_k \sigma_k$.

Discretization Accuracy. Let c_k be the difference between the true moment μ_k and the approximate moment $\hat{\mu}_k$, obtained by means of a discretization. It will be useful to normalize this difference by the standard deviation of the moment, which may be obtained with Eq. 12. We write the accuracy of the approximation as

$$\delta_k = \frac{c_k}{\sqrt{m_k^2}} = \frac{\mu_k - \hat{\mu}_k}{\sigma_k}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (14)$$

Computing S-Equivalence. Now, suppose that instead of using a discretization, we estimate μ_k by means of MC simulation. Let the estimated value, which is an average, given S samples be $\bar{\mu}_{k;S}$. According to the *central limit theorem* (CLT), $\bar{\mu}_{k;S}$ is normally distributed with mean μ_k and variance σ_k^2 / S for large S. Thus, the probability that the simulation will more accurately estimate μ_k than the discretization method is (see Appendix A for derivation) given by

$$P\left(\left|\bar{\mu}_{k;S} - \mu_k\right| \le c_k\right) = 2\Phi(x) - 1, \quad \ldots\ldots\ldots\ldots\ldots\ldots (15)$$

where $x = \delta_k \sqrt{S}$ and Φ is the standard normal CDF. If we want this probability to be τ, then we must take

$$S = \delta_k^{-2} \Phi^{-1}\left(\frac{\tau + 1}{2}\right)^2. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (16)$$

samples (see Appendix A for derivation). For example, if we want the simulation to have a 95% chance of estimating the kth raw moment more accurately than the discretization, we must take $S = \delta_k^{-2} \Phi^{-1}(1.95/2)^2 \approx \delta_k^{-2} \cdot 1.96^2$ samples. Thus, Eq. 16 establishes an equivalence between discretization and simulation (Pfeifer et al. 1991). [Pfeifer et. al (1991) determined S-equivalence for matching the mean. They did not consider higher moments or use the Edgeworth expansion, as we do here.]

Eq. 16 will not work well when the underlying random variable X^k is highly skewed or kurtotic. In these cases, the CLT approximation may not be accurate for the number of trials that Eq. 16 says we need to perform. The CLT holds under quite general conditions and states that the sum of an infinite number of random variables converges in distribution to the normal. However, it does not specify how quickly this convergence will take place. In this case, we must rely on the Edgeworth expansion (Cramér 1946; Hall 1992), which corrects the CLT for higher underlying moments. For the distributions we consider here, this issue primarily affects the log-normal distribution and, to a lesser degree, the exponential. In this case, the probability that the MC simulation will estimate μ_k more accurately than the discretization method is (see Appendix A for derivation) given by

$$P\left(\left|\bar{\mu}_{k;S} - \mu_k\right| \le c_k\right) = 2\Phi(x) - 1 - \frac{1}{S}\left[\begin{array}{l} \frac{1}{12}\gamma_4^k\left(x^3 - 3x\right)\phi(x) \\ + \frac{1}{36}\left(\gamma_3^k\right)^2\left(x^5 - 10x^3 + 15x\right)\phi(x) \end{array}\right] + O\left(S^{-2}\right), \quad \ldots\ldots (17)$$

TABLE 7—95% *S*-EQUIVALENCES FOR DISCRETIZATION SHORTCUTS

Raw	EPT					GQN				
Moment	$U(0,b)$	$N(0,\sigma)$	$T(0,b,b/2)$	$E(\lambda)$	$L(\mu,1)$	$U(0,b)$	$N(0,\sigma)$	$T(0,b,b/2)$	$E(\lambda)$	$L(\mu,1)$
First (mean)	∞	∞	∞	>1MM*		∞	∞	∞	>1MM	161,943
Second	4,830	>1MM	67,654	57,470	2,781	1,930	∞	644,360	>1MM	4,583
Third	1,940	∞	24,388	3,918	48,591	775	∞	231,941	14,592	56,102
Raw	ESM					MCS				
Moment	$U(0,b)$	$N(0,\sigma)$	$T(0,b,b/2)$	$E(\lambda)$	$L(\mu,1)$	$U(0,b)$	$N(0,\sigma)$	$T(0,b,b/2)$	$E(\lambda)$	$L(\mu,1)$
First (mean)	∞	∞	∞	>1MM	2,495	∞	∞	∞	1,451	560
Second	2,128	36,165	9,745	1,674	941	30,732	240	14,068	407	676
Third	855	∞	3,508	498	34,475	12,347	∞	5,064	207	35,520
Raw	MRO					ZDT				
Moment	$U(0,b)$	$N(0,\sigma)$	$T(0,b,b/2)$	$E(\lambda)$	$L(\mu,1)$	$U(0,b)$	$N(0,\sigma)$	$T(0,b,b/2)$	$E(\lambda)$	$L(\mu,1)$
First (mean)	∞	∞	∞	13,529	1,702	∞	∞	∞	120,295	2,722
Second	77,873	1,840	115,500	1,264	988	1,046	904,694	5,986	1,716	895
Third	31,293	∞	41,584	554	35,758	420	∞	2,155	448	33,688

*Note: >1MM = more than 1 million.

where $x = \delta_k \sqrt{S}$, as before, and ϕ is the standard normal PDF. The term $O(S^{-2})$ signifies that the remaining terms are at most of order S^{-2}. The terms $\gamma_3^k = m_k^3 m_k^{-3/2} = E[(X^k - \mu_k)^3]E[(X^k - \mu_k)^2]^{-3/2}$ and $\gamma_4^k = m_k^4 m_k^{-2} = E[(X^k - \mu_k)^4]E[(X^k - \mu_k)^2]^{-2} - 3$ are the skewness and kurtosis, respectively. If the skewness and kurtosis are zero (or small), or if S is very large, then Eq. 17 reduces to Eq. 15 and we can use Eq. 16 to find the equivalent number of samples. On the other hand, if the skewness and kurtosis are nonnegligible, then we must numerically solve Eq. 17 for S.

Lerche and Mudford (2005a, 2005b) investigate the number of samples required to estimate the mean of several distributions including log-normal and exponential and base their estimates on the CLT. This approximation, while not fully appropriate, should not have induced substantial errors in their case because of their focus on estimating only the mean. However, we are interested in estimating higher moments, and these higher moments (e.g., $E[X^4]$) will induce additional skewness and kurtosis and the CLT will fail to work well in these cases. For example, for the log-normal distribution that we discuss here, we find that Eq. 16 overestimates the required number of samples by almost 25%. In the case of the log-normal, negative values are not possible. This causes Eq. 16, which is based on the CLT and assumes that the mean is normally distributed, to require more samples so as to reduce the variance of the mean enough such that negative values are very unlikely—at least less than (1 – 0.95)/2 = 2.5%. The Edgeworth expansion performs better by taking the kurtosis and the skewness into account. These normalized central moments are *very* large in the case of the log-normal—especially when we are considering the distribution of X^3 and X^4.

Table 7 provides the 95% *S*-equivalences based on Eq. 17 for the EPT, GQN, ESM, MCS, MRO, and ZDT shortcuts for the uniform $U(0, b)$, normal $N(0, \sigma)$, triangular $T(0, b, b/2)$, exponential $E(\lambda)$, and log-normal $L(\mu, 1)$ distributions. These results hold for any (see Appendix A)

- Exponential distribution
- Uniform distribution bounded by zero on one side, $U(0, b)$
- Normal distribution centered at zero, $N(0, \sigma)$
- Symmetric triangular distribution bounded by zero on one side, $T(0, b, b/2)$
- Log-normal distribution with unit variance (of ln X), $L(\mu, 1)$

While the results in Table 7 do not hold for all possible distributions within a given family, they should give the reader a sense for the magnitude of MC samples that are required to match a particular discretization.

These results are striking. First, setting aside the fact that all the approximations we consider perfectly match the mean of symmetric distributions [U, N, $T(a, b, b/2)$], most of the shortcuts are equivalent to thousands, if not tens of thousands, of MC samples. For example, it would take 57,469 MC samples to have a 95% chance of estimating the second raw moment of an exponential with greater accuracy than EPT, which Pearson and Tukey (1965) did not even suggest using to estimate any moment beyond the mean. Second, EPT and GQN are clearly dominant, having larger *S*-equivalences than the other approximations. These approximations do very well on everything except the third moment of uniform distributions. Consider estimating the mean of log-normal distribution: EPT and GQN are equivalent to almost 12 and 65 times more samples, respectively, than ESM (29,499 or 161,943 compared to 2,495). Third, these results serve to emphasize our earlier conclusion that the performance of ESM and MCS is quite poor in some cases. For example, ESM and MCS are only equivalent to 941 and 676 MC samples, respectively, in terms of estimating the second moment of the $L(\mu, 1)$ distribution. (The discretization shortcuts do not provide unbiased estimates of the underlying moments. We do not address the issue of bias here. Rather, we assume that over- and underestimates of the moments are equally costly.) While the direction of this result is not surprising, given that we know these methods underestimate variance, its magnitude is striking. Use of ESM to model the uncertainty of a log-normal distribution, which occurs whenever it is applied to estimations of oil or gas reserves, is equivalent to running fewer than 1,000 MC trials!

A large *S*-equivalence is a result of two possible factors: (1) high accuracy of the discretization method and (2) the difficulty of simulating the underlying random variable. As a case in point, consider the EPT and GQN approximations of the log-normal distribution. The *S*-equivalences for estimating the mean are large because the discretizations estimate the mean closely. The *S*-equivalences for the third moment are large because approximating the third moment of a log-normal distribution by means of MC simulation is very difficult, requiring many tens of thousands of MC samples. Given the large number of samples that are required in most cases, we see that discretization is a viable, and in some cases preferable, alternative to MC simulation. Of course, we are interested in preserving the moments of the output (e.g., NPV), which might be a function of many input uncertainties and several downstream decisions. We do not address this more complicated issue here.

Does Discretization Matter?

"New" methodologies are often resisted on the grounds that they will not obviously make a material difference. (We place "new" in quotes because Pearson-Tukey predates Swanson-Megill by almost a decade and Gaussian quadrature predates all other methods by

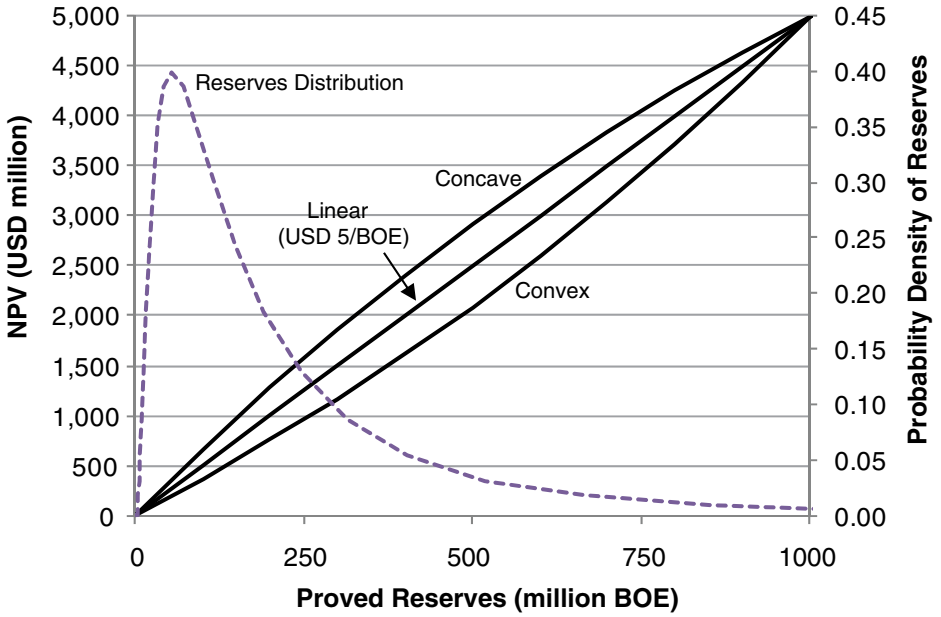

Fig. 5—Illustrative example used to show inaccuracy of various approximations.

at least a century.) While we do not have space to fully address this issue here, we note that most of the methods presented in this paper are no more complicated than existing approximations, but are more accurate. For example,

- EPT and GQN are three-point approximations, like ESM and MCS, but are more accurate across a range of distributions.
- The approximations given in Tables 3 through 5 are three-point approximations that match the first three moments of the listed distributions—again, exceeding the performance of ESM and MCS.
- The four-point approximations given in Table 6 match the mean and variance of log-normal distributions, resulting in better performance than the three-point ESM and MCS shortcuts.
- The log-SM method that we outline will perfectly match all the moments of the log-normal and is only slightly more complex than SM.
- Other methods that we discuss, such as moment matching for distributions not listed in Tables 4 through 7, are more complex than current practice. However, this increase in complexity will increase the accuracy of the results (much in the same way that 3D reservoir simulation is more complex, but more accurate, than 2D simulation).

However, we realize that some will require convincing before changing their practice. To address this, we offer the following, necessarily simple, example.

TABLE 8—APPROXIMATION ERROR IN MEAN FOR ILLUSTRATIVE EXAMPLE

Approximation	Concave	Linear	Convex
4-pt Lognormal	–1%	0%	0%
EPT	0%	–1%	–4%
GQN	0%	–1%	–3%
ESM	–3%	–5%	–9%
MCS	–8%	–11%	–15%
MRO	–4%	–6%	–10%
ZDT	–2%	–5%	–9%
3-pt Bracket Median	–14%	–17%	–21%
4-pt Bracket Median	–10%	–12%	–16%
3-pt Bracket Mean	2%	0%	–4%
4-pt Bracket Mean	1%	0%	–2%

Illustrative Example. Suppose an oil company is considering the purchase of a prospect that contains an uncertain volume of oil. For the sake of argument, suppose the reserves are believed to be log-normally distributed with a mean of 90 million BOE and a standard deviation of 118 million BOE. These parameters correspond to a standard deviation for log reserves of 1.0, and, thus, we can make use of the four-point approximation given in Table 6. To illustrate the accuracy of each approximation method, we will consider three different valuation scenarios. In the first, the company believes, on the basis of market transactions (Howard and Harp 2009), that proved reserves in this geographic location and depositional environment are worth USD 5 per BOE. This case is labeled "Linear" in **Fig. 5.** We also consider a case where the value function is convex and another that is concave. The concave case could correspond to a situation where the host government takes a small share of small fields to increase the probability of development but will take a proportionally larger share as field size increases. The convex case might represent large initial fixed costs, perhaps for infrastructure, that do not scale in proportion to field size. Whatever the case may be, the important point, as far as the example is concerned, is that discretization accuracy depends upon the shape of the value function.

We next apply the discretization methods discussed in this paper to the reserves distribution and compute the expected NPV of the prospect. For the bracket-median and bracket-mean discretizations, we consider three- and four-point approximations with weightings of 25-50-25 and 10-40-40-10, respectively. We compare the approximate value from each discretization to the exact value, which we obtain by means of 100,000 MC samples. The error for each approximation is presented in **Table 8.**

The errors in the linear case are simply the errors of each discretization in estimating mean reserves. We see that the four-point log-normal approximation and the EPT, GQN, and the bracket-mean approaches estimate the expected NPV to within 1%. The other methods underestimate the mean by at least 5%; the bracket-median approach is especially poor. Errors in the concave and convex cases differ from the linear case because in these situations the variance (and other moments) of the input distribution is important. Because most of discretization methods misestimate the variance, they will misestimate the expected value of the output (NPV). The errors are less in the concave case because the concavity of the value function serves to reduce the impact of misestimating the variance of the input distribution. In the convex case, we see that the four-point log-normal and the four-point bracket mean perform very well. **Table 9** presents the error in the variance of NPV. In

TABLE 9—APPROXIMATION ERROR IN VARIANCE FOR ILLUSTRATIVE EXAMPLE

Approximation	Concave	Linear	Convex
4-pt Lognormal	30%	0%	–57%
EPT	–18%	–36%	–72%
GQN	–11%	–30%	–69%
ESM	–45%	–60%	–83%
MCS	–53%	–66%	–86%
MRO	–42%	–57%	–82%
ZDT	–47%	–62%	–84%
3-pt Bracket Median	–66%	–75%	–90%
4-pt Bracket Median	–48%	–60%	–83%
3-pt Bracket Mean	–38%	–54%	–80%
4-pt Bracket Mean	–21%	–37%	–71%

this case, we see that all the approximations underestimate the output variance. The performance of ESM, MCS, and the bracket-median approaches are especially poor, while the four-point log-normal and the four-point bracket-mean discretizations perform significantly better. Of course, the four-point log-normal and the bracket-mean approximations have been specifically tailored to the underlying distribution, which demonstrates the importance of this practice.

Is underestimating the mean by 10% and the variance by 50% a problem? Clearly, if the company is setting a purchase (or sale) price, underestimating the value of the prospect by 10% could undermine the opportunity. This would be especially true in a competitive situation, such as bidding. In addition, the optimal bid amount depends critically upon the company's estimate of uncertainty. Underestimating the variance by 50–80% may increase the probability of overpaying for the property. Thus, we see that discretization matters and could have a material impact on decision making. In what other part of the business would misestimating key performance metrics by 10–80% be acceptable?

Conclusion: Recommendations and Discussion

We conclude with a set of recommendations and observations regarding the oil and gas industry's approach to probabilistic modeling.

Recommendations. If one is interested in estimating the moments of an output distribution, then closely matching the moments of the input distributions is a necessary requirement. In this case, the best discretization methods presented in this paper are very accurate, matching $2N$ moments (including the zeroth) with only N points. They could, therefore, be used instead of MC simulation, at least early in an analysis, to quantify uncertainty with only a few evaluations of the output function. The primary recommendations that follow from our work are

- For maximum accuracy, we should use moment matching and apply it directly to each input distribution. For several distribution families, the appropriate weights and values have already been calculated (see Table 1 and Tables 3 through 7). In other cases, the quadrature could be calculated on a case-by-case basis using methods detailed in Stroud and Secrest (1966), Miller and Rice (1983), Smith (1993), or Davis and Rabinowitz (1984), for example.
- If moment matching is too difficult to implement or communicate then the equal-areas approach is not an unreasonable alternative. However, one must bear in mind that this approach will tend to underestimate the variance (and other higher moments).
- If one wishes to fix the values at the P90, P50, P10 or weights at 25-50-25 or 30-40-30 for communication, assessment, or computational reasons, then using the moment-matching discretizations in Tables 3 through 5 will result in maximum accuracy.
- When dealing with log-normal distributions, work with the log of the random variable, which translates it to a normal random variable. Then, apply the discretization methods to the transformed variable.
- SM (or ESM), MCS, and ZDT should not be used as part of a final analysis (recall that McNamee and Celona explicitly warned against using their shortcut in this way). Direct application of ESM to log-normal variables, as is common, to estimate their moments (e.g., mean, variance) should be considered an unacceptable professional practice.
- Well-chosen discrete approximations are equivalent to tens of thousands of MC samples. This argues for greater use of discretization methods.

Discussion and Conclusion. Despite attempts to justify it (Hurst et al. 2000; Rose 2001), SM has no theoretical justification for use with any distribution other than normal (direct use with the log-normal is especially error prone). Megill noted its problems nearly 40 years ago. Why would we expect MCS or ESM/SM (or the other shortcuts), which are *symmetric,* to preserve the mean of skewed distributions? [Some analysts have told us that MCS (25-50-25) is good for symmetric distributions, but that ESM (30-40-30) should be used when the distribution is skewed!] Why have we institutionalized a method known to be biased? Megill's answer was that SM offers protection from uncertainty because its estimates are known to be biased low (Megill 1984). This is in much the same spirit as using a high discount rate or a low oil price to account for uncertainty. We do not support such adjustments; decision makers should be provided with unbiased estimates of the risks facing the company.

Another explanation for the acceptance of SM is the industry's focus on the mean and "risking" a prospect (Bickel and Bratvold 2008), to the exclusion of other moments. Paraphrasing Steven Jay Gould, "the mean is not the message" (Gould 1985). Risking a prospect does nothing either to understand the risk or to help manage it. It simply implies that the mean, the probability-weighted average, was calculated. Decision makers are not indifferent to all projects with the same mean. Rather, they want to understand the surrounding uncertainty and the risk. Furthermore, as demonstrated by Eq. 5 and illustrated in the preceding section, one may not accurately determine the mean of an output distribution without accurately representing the mean and higher moments of the input distributions.

Why do we continue to use ESM when more-accurate alternatives exist? Why do we spend millions of dollars on reservoir-simulation models and then represent the output of those models in a decision tree with simplistic discretizations? Perhaps it is simply a matter of tradition and path dependence: "We have always done it this way," "That is the way everybody else does it," or "That is the way I was taught." In fact, some companies mandate that either MCS or ESM be used in project valuations. We hope that our paper will encourage improved practice.

Nomenclature

c_k = difference between true moment and approximate moment
$E[-]$ = expectation operator
$E(\lambda)$ = exponential random variable with a mean of $1\ \lambda$
f = PDF
F = CDF
G = excess distribution function
$L(\mu, \sigma)$ = log-normal random variable with mean of $\ln X$ of μ and standard deviation of $\ln X$ of σ
$m_k = E[(X - \mu_l)^k]$ = kth central moment of X
N = number of discretization points
$N(\mu, \sigma)$ = normal random variable with mean of μ and standard deviation of σ
O = order of approximation
p = probability
$P(x)$ = polynomial in x
PZ = Zth percentile of CDF

S = number of MC samples
$T(a, b, c)$ = triangular random variable between a and b with a mode at c
$U(a, b)$ = uniform random variable between a and b
v = value function
x = realization of random variable X
X = random variable
α_i = percentile
γ_3 = skewness
γ_4 = kurtosis
δ_k = accuracy of discrete approximation of kth moment
$\hat{\mu}_k$ = kth moment obtained from discrete approximation
$\bar{\mu}_{k;S}$ = kth moment obtained from average of S MC samples
$\mu_k = E[X^k]$ = k-th raw moment of X
$\sigma^2 = m_2$ = variance
Φ = standard normal CDF

Acknowledgments

Eric Bickel thanks the Research Partnership to Secure Energy for America (Grant #07122-35) for supporting this research. Eric Bickel and Larry Lake, the W.A. (Monty) Moncrief Centennial Chair in the Department of Petroleum and Geosystems Engineering, both thank the Center for Petroleum Asset Risk Management at the University of Texas at Austin for its support. The authors thank David Morton for his comments on a draft of this paper.

References

Arild, Ø., Lohne, H.P., and Bratvold, R. 2008. A Monte Carlo Approach to Value of Information. Paper SPE IPTC 11969 presented at the International Petroleum Technology Conference, Kuala Lumpur, 3–5 December. http://dx.doi.org/10.2523/11969-MS.

Bickel, J.E. and Bratvold, R.B. 2008. From Uncertainty Quantification to Decision Making in the Oil and Gas Industry. *Energy Exploration and Exploitation* **26** (5): 311–325. http://dx.doi.org/10.1260/014459808787945344.

Bickel, J.E., Gibson, R.L., McVay, D.A., Pickering, S., and Waggoner, J. 2008. Quantifying the Reliability and Value 3D Land Seismic. *SPE Res Eval & Eng* **11** (5) 832–841. SPE-102340-PA. http://dx.doi.org/10.2118/102340-PA.

Clemen, R.T. and Reilly, T. 2001. *Making Hard Decisions*, second edition. Pacific Grove, California: Duxbury Press.

Cramér, H. 1946. *Mathematical Methods in Statistics*. Princeton, New Jersey: Princeton University Press.

Davis, P.J. and Rabinowitz, P. 1984. *Methods of Numerical Integration*, second edition. Orlando, Florida: Computer Science and Applied Mathematics, Academic Press.

D'Errico, J.R. and Zaino, N.A. 1988. Statistical Tolerancing Using a Modification of Taguchi's Method. *Technometrics* **30** (4): 397–405.

Gauss, C.F. 1866. Methodus nova integralium valores per approximationen inveniendi. In *Carl Friedrich Gauss Werke,* 163–196. Göttingen, Germany: Königlichen Gesellschaft der Wissenschaften.

Gould, S.J. 1985. The Median Isn't the Message. *Discover* (June): 40–42.

Hall, P. 1992. *The Bootstrap and Edgeworth Expansion*. New York: Springer Series in Statistics, Springer-Verlag.

Howard, A.W. and Harp, A.B. Jr. 2009. Oil and Gas Company Valuations. *Business Valuation Review* **28** (1): 30–35.

Hurst, A., Brown, G.C., and Swanson, R.I. 2000. Swanson's 30-40-30 rule. *AAPG Bulletin* **84** (12): 1883–1891.

Keefer, D.L. 1994. Certainty Equivalents for Three-Point Discrete-Distribution Approximations. *Management Science* **40** (6): 760–773. http://dx.doi.org/10.1287/mnsc.40.6.760.

Keefer, D.L. and Bodily, S.E. 1983. Three-Point Approximations for Continuous Random Variables. *Management Science* **29** (5): 595–609. http://dx.doi.org/10.1287/mnsc.29.5.595.

Keefer, D.L. and Pollock, S.M. 1980. Approximations and Sensitivity in Multiobjective Resource Allocation. *Operations Research* **28** (1): 114–128. http://dx.doi.org/10.1287/opre.28.1.114.

Lerche, I. and Mudford, B.S. 2005a. How Many Monte Carlo Simulations Does One Need to do? *Energy Exploration & Exploitation* **23** (6): 405–427. http://dx.doi.org/10.1260/014459805776986876.

Lerche, I. and Mudford, B.S. 2005b. How Many Monte Carlo Simulations Does One Need to do? II. Lognormal, Binomial, Cauchy, and Exponential Distributions. *Energy Exploration & Exploitation* **23** (6): 429–461. http://dx.doi.org/10.1260/014459805776986894.

McNamee, P. and Celona, J. 1990. *Decision Analysis with Supertree*, second edition. San Francisco, California: Scientific Press.

McNamee, P. and Celona, J. 1991. *Decision Analysis with Supertree: Instructor's Manual*, second edition. San Francisco, CA: Scientific Press.

Megill, R.E. 1984. *An Introduction to Risk Analysis*, second edition. Tulsa, Oklahoma: PennWell.

Miller, A.C. III and Rice, T.R. 1983. Discrete Approximations of Probability Distributions. *Management Science* **29** (3): 352–362. http://dx.doi.org/10.1287/mnsc.29.3.352.

Papoulis, A. 1984. *Probability, Random Variables, and Stochastic Processes*, second edition. New York: Series in Electrical Engineering, McGraw-Hill.

Pearson, E.S. and Tukey, J.W. 1965. Approximate means and standard deviations based on distances between percentage points of frequency curves. *Biometrika* **52** (3–4): 553–546. http://dx.doi.org/10.1093/biomet/52.3-4.533.

Pfeifer, P.E., Bodily, S.E., and Frey, S.C. Jr. 1991. Pearson-Tukey Three-Point Approximations Versus Monte-Carlo Simulation. *Decision Sciences* **22** (1): 74–90. http://dx.doi.org/10.1111/j.1540-5915.1991.tb01263.x.

Prange, M., Bailey, W., and Kuchuk, F. 2009. Quantifying the Value of a Future Well Test with Measurement Uncertainty. Paper IPTC 13595 presented at the International Petroleum Technology Conference, Doha, Qatar, 7–9 December. http://dx.doi.org/10.2523/13595-MS.

Rose, P.R. 2001. *Risk Analysis and Management of Petroleum Exploration Ventures*. Tulsa, Oklahoma: Methods in Exploration Series, AAPG.

Smith, J.E. 1993. Moment Methods for Decision Analysis. *Management Science* **39** (3): 340–358. http://dx.doi.org/10.1287/mnsc.39.3.340.

Stibolt, R.D. and Lehman, J. 1993. The Value of a Seismic Option. Paper SPE 25821 presented at the SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, 29–30 March. http://dx.doi.org/10.2118/25821-MS.

Stroud, A.H. and Secrest, D. 1966. *Gaussian Quadrature Formulas*. Englewood Cliffs, New Jersey: Prentice-Hall Series in Automatic Computation, Prentice-Hall.

Taguchi, G. 1978. Performance Analysis Design. *International Journal of Production Research* **16** (6): 521–530. http://dx.doi.org/10.1080/00207547808930043.

Willigers, B.J.A. 2009. Practical Portfolio Simulation: Determining the Precision of a Probability Distribution of a Large Asset Portfolio when the Underlying Project Economics are Captured by a Small Number of Discrete Scenarios. Paper SPE 124180 presented at the SPE Annual Technical Conference and Exhibition, New Orleans, 4–7 October. http://dx.doi.org/10.2118/124180-MS.

Zaino, N.A. and D'Errico, J. 1989. Optimal Discrete Approximations for Continuous Outcomes with Applications in Decision and Risk Analysis. *Journal of the Operational Research Society* **40** (4): 379–388.

Appendix A

Probability the MC Simulation Will Be More Accurate Than the Approximation.

$$
\begin{aligned}
P\left(\left|\bar{\mu}_{k;S}-\mu_k\right|\le c\right) &= P\left(\bar{\mu}_{k;S}-\mu_k\le +c\right)-P\left(\bar{\mu}_{k;S}-\mu_k\le -c\right)\\
&= P\left(\frac{\bar{\mu}_{k;S}-\mu_k}{\sigma_k/\sqrt{S}}<\frac{+c}{\sigma_k/\sqrt{S}}\right)-P\left(\frac{\bar{\mu}_{k;S}-\mu_k}{\sigma_k/\sqrt{S}}<\frac{-c}{\sigma_k/\sqrt{S}}\right)\\
&= P\left(\frac{\bar{\mu}_{k;S}-\mu_k}{\sigma_k/\sqrt{S}}<\delta\sqrt{S}\right)-P\left(\frac{\bar{\mu}_{k;S}-\mu_k}{\sigma_k/\sqrt{S}}<-\delta\sqrt{S}\right)\\
&= \Phi\left(\delta\sqrt{S}\right)-\Phi\left(-\delta\sqrt{S}\right)=\Phi\left(\delta\sqrt{S}\right)-\left[1-\Phi\left(\delta\sqrt{S}\right)\right]\\
&= 2\Phi\left(\delta\sqrt{S}\right)-1. \qquad \text{.....(A-1)}
\end{aligned}
$$

S-Equivalence for CLT.

$$2\Phi\left(\delta\sqrt{S}\right)-1=\tau$$
$$\delta\sqrt{S}=\Phi^{-1}\left(\frac{\tau+1}{2}\right). \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(A-2)}$$
$$S=\delta^{-2}\Phi^{-1}\left(\frac{\tau+1}{2}\right)^2$$

Edgeworth Expansion. The first three terms of the Edgeworth expansion are

$$F_S(x)=\Phi(x)-\frac{1}{S^{1/2}}\left(\frac{1}{6}\frac{\kappa_3}{\kappa_2^{3/2}}\Phi^{(3)}(x)\right)+\frac{1}{S}\left(\begin{array}{l}\frac{1}{24}\frac{\kappa_4}{\kappa_2^{4/2}}\Phi^{(4)}(x)\\+\frac{1}{72}\left(\frac{\kappa_3}{\kappa_2^{3/2}}\right)^2\Phi^{(6)}(x)\end{array}\right)-\frac{1}{S^{3/2}}\left(\begin{array}{l}\frac{1}{120}\frac{\kappa_5}{\kappa_2^{5/2}}\Phi^{(5)}(x)+\frac{1}{144}\frac{\kappa_3}{\kappa_2^{3/2}}\frac{\kappa_4}{\kappa_2^{4/2}}\Phi^{(7)}(x)\\+\frac{1}{1296}\left(\frac{\kappa_3}{\kappa_2^{3/2}}\right)^3\Phi^{(9)}(x)\end{array}\right)+O\left(S^{-5/2}\right), \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(A-3)}$$

where κ_r is the rth cumulant, κ_2 is the variance, and κ_1 is the mean. $\Phi^{(j)}$ is the jth derivative of $\Phi^{(j)}$. Thus, $\Phi^{(j)}=\phi^{(j-1)}$, where ϕ is the standard normal PDF. The jth derivative of the standard normal PDF is $\phi^{(j)}=(-1)^j H_j(x)\phi(x)$, where $H_j(x)$ is the Hermite polynomial of order j. Thus, $\Phi^{(j)}=\phi^{(j-1)}=(-1)^{j-1}H_{j-1}(x)\phi(x)$. The Hermite polynomials are an even (odd) function when $j-1$ is even (odd). Thus, we have

$$F_S(x)=\Phi(x)-\frac{1}{S^{1/2}}\left(\frac{1}{6}\frac{\kappa_3}{\kappa_2^{3/2}}H_2(x)\phi(x)\right)-\frac{1}{S}\left(\begin{array}{l}\frac{1}{24}\frac{\kappa_4}{\kappa_2^{4/2}}H_3(x)\phi(x)\\+\frac{1}{72}\left(\frac{\kappa_3}{\kappa_2^{3/2}}\right)^2H_5(x)\phi(x)\end{array}\right)-\frac{1}{S^{3/2}}\left(\begin{array}{l}\frac{1}{120}\frac{\kappa_5}{\kappa_2^{5/2}}H_4(x)\phi(x)\\+\frac{1}{144}\frac{\kappa_3}{\kappa_2^{3/2}}\frac{\kappa_4}{\kappa_2^{4/2}}H_6(x)\phi(x)\\+\frac{1}{1296}\left(\frac{\kappa_3}{\kappa_2^{3/2}}\right)^3H_8(x)\phi(x)\end{array}\right)+O\left(S^{-2}\right). \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(A-4)}$$

In the discretization work, we are looking at the probability of being within the interval $[-x, x]$, which is $F_S(x) - F_S(-x)$. Because H_j is even when j is even and $\phi(x)$ is an odd function, the terms of order $S^{-j/2}$ will cancel in the subtraction when j is odd and we will be left with

$$F_S(x)-F_S(-x)=2\Phi(x)-1-\frac{1}{S}\left[\begin{array}{l}\frac{1}{12}\frac{\kappa_4}{\kappa_2^{4/2}}H_3(x)\phi(x)\\+\frac{1}{36}\left(\frac{\kappa_3}{\kappa_2^{3/2}}\right)^2H_5(x)\phi(x)\end{array}\right]+O\left(S^{-2}\right)$$
$$=2\Phi(x)-1-\frac{1}{S}\left[\begin{array}{l}\frac{1}{12}\gamma_4H_3(x)\phi(x)\\+\frac{1}{36}\gamma_3^2H_5(x)\phi(x)\end{array}\right]+O\left(S^{-2}\right), \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(A-5)}$$

where γ_3 is the skewness and γ_4 is the kurtosis. The third and fifth Hermite polynomials are $H_3(x)=x^3-3x$ and $H_5(x)=x^5-10x^3+15x$ Substituting these values into Eq. A-5 yields Eq. 17.

Exponential S-Equivalence. The inverse CDF for the exponential is $F^{-1}(\alpha_i)=-\lambda^{-1}\ln(1-\alpha_i)$. The approximate moment for approximation A is then

$$\hat{\mu}_k=\sum_{i=1}^N p_i^A\left(-\lambda^{-1}\ln(1-\alpha_i^A)\right)^k=-\lambda^{-k}\sum_{i=1}^N p_i^A\left(\ln(1-\alpha_i^A)\right)^k. \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(A-6)}$$

The kth raw moment of the exponential is

$$\mu_k=-\lambda^{-k}\Gamma(k+1). \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(A-7)}$$

The variance of the kth raw moment is

$$\sigma_k^2=-\lambda^{-2k}\left(\Gamma(2k+1)-\Gamma(k+1)^2\right). \quad \ldots\ldots\ldots\ldots\ldots\text{(A-8)}$$

Thus, we then have

$$\delta_k=\frac{\mu_k-\hat{\mu}_k}{\sigma_k}=\frac{\lambda^{-k}\Gamma(k+1)-\lambda^{-k}\sum_{i=1}^N p_i^A\left(\ln(1-u_i^A)\right)^k}{\lambda^{-k}\left(\Gamma(2k+1)-\Gamma(k+1)^2\right)^{1/2}}=\frac{\Gamma(k+1)-\sum_{i=1}^N p_i^A\left(\ln(1-u_i^A)\right)^k}{\left(\Gamma(2k+1)-\Gamma(k+1)^2\right)^{1/2}}. \quad \ldots\ldots\ldots\ldots\ldots\text{(A-9)}$$

We see that δ is independent of λ and that the S-equivalences that we list in Table 7 hold for any exponential PDF.

Uniform *S*-Equivalence. The inverse CDF for the uniform is $F^{-1}(\alpha_i)=a+\alpha_i(b-a)$. The approximate kth raw moment for approximation A is then

$$\hat{\mu}_k=\sum_{i=1}^N p_i^A\left(a+\alpha_i^A(b-a)\right)^k. \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(A-10)}$$

The true kth raw moment of the uniform is

$$\mu_k=\frac{a^{k+1}-b^{k+1}}{(a-b)(k+1)}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(A-11)}$$

The variance of the kth raw moment is

$$\sigma_k^2=\frac{\left(a^{2k+2}+b^{2k+2}\right)k^2-a^{2k+1}b(k+1)^2+ab\left[2(ab)^k(2k+1)-b^{2k}(k+1)^2\right]}{(a-b)^2(k+1)^2(2k+1)}. \quad \ldots\ldots\ldots\ldots\text{(A-12)}$$

Thus, we have

$$\delta_k=\frac{\mu_k-\hat{\mu}_k}{\sigma_k}=\frac{\frac{a^{k+1}-b^{k+1}}{(a-b)(k+1)}-\sum_{i=1}^N p_i^A\left(a+u_i^A(b-a)\right)^k}{\left\{\frac{\left(a^{2k+2}+b^{2k+2}\right)k^2-a^{2k+1}b(k+1)^2+ab\left[2(ab)^k(2k+1)-b^{2k}(k+1)^2\right]}{(a-b)^2(k+1)^2(2k+1)}\right\}^{1/2}}. \quad \ldots\ldots\ldots\ldots\ldots\text{(A-13)}$$

If $a = 0$, then we have

$$\delta_k=\frac{\mu_k-\hat{\mu}_k}{\sigma_k}=\frac{\left(\frac{1}{k+1}-\sum_{i=1}^N p_i^A\left(u_i^A\right)^k\right)}{\frac{k}{k+1}\frac{1}{(2k+1)^{1/2}}}, \quad \ldots\ldots\ldots\ldots\text{(A-14)}$$

and we see that δ is independent of b and that the S-equivalences that we list in Table 7 hold for any $U(0,b)$. If a is nonzero, then S-equivalences will differ from those shown in Table 7.

Normal *S*-Equivalence. The inverse CDF for the normal is $F^{-1}(\alpha_i) = \mu + \sigma\Phi^{-1}(\alpha_i)$. The approximate kth raw moment for approximation A is then

$$\hat{\mu}_k = \sum_{i=1}^{N} p_i^A \left(\mu + \sigma\Phi^{-1}(\alpha_i^A)\right)^k. \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(A-15)}$$

In the case of the normal, the raw moments are quite complex and we restrict our attention to the case where $\mu = 0$. In this case, the true kth raw moment of the normal is

$$\mu_k = \pi^{-1/2}\sigma^k 2^{k/2-1}\left(1+e^{ik\pi}\right)\Gamma\left(\frac{k+1}{2}\right). \quad \ldots\ldots\ldots\ldots\text{(A-16)}$$

Similarly, the variance of the kth raw moment is

$$\sigma_k^2 = \pi^{-1}2^{k-2}\sigma^{2k}\begin{bmatrix} 2\left(1+e^{2ik\pi}\right)\sqrt{\pi}\Gamma\left(k+\frac{1}{2}\right) \\ -\left(1+e^{ik\pi}\right)^2\Gamma\left(\frac{k+1}{2}\right)^2 \end{bmatrix}. \quad \ldots\ldots\ldots\text{(A-17)}$$

Thus, we have

$$\delta_k = \frac{\mu_k - \hat{\mu}_k}{\sigma_k} = \frac{\begin{matrix} 2^{k/2-1}\left(1+e^{ik\pi}\right)\Gamma\left(\frac{k+1}{2}\right) \\ -\sqrt{\pi}\sum_{i=1}^{N} p_i^A \left(\Phi^{-1}(u_i^A)\right)^k \end{matrix}}{\left\{2^{k-2}\begin{bmatrix} 2\left(1+e^{2ik\pi}\right)\sqrt{\pi}\Gamma\left(k+\frac{1}{2}\right) \\ -\left(1+e^{ik\pi}\right)^2\Gamma\left(\frac{k+1}{2}\right)^2 \end{bmatrix}\right\}^{1/2}}, \quad \ldots\text{(A-18)}$$

and we see that δ is independent of σ and that the *S*-equivalences that we list in Table 7 hold for any $N(0, \sigma)$.

Triangular *S*-Equivalence. The inverse CDF for the triangular is

$$F^{-1}(\alpha_i) = \begin{cases} a + \sqrt{(b-a)(c-a)\alpha_i} & 0 \le \alpha_i < F(c) \\ b - \sqrt{(b-a)(b-c)(1-\alpha_i)} & F(c) \le \alpha_i \le 1 \end{cases}. \quad \ldots\text{(A-19)}$$

The approximate kth raw moment for approximation A is then

$$\hat{\mu}_k = \begin{matrix} \sum_{i=1}^{N_c} p_i^A \left(a + \sqrt{(b-a)(c-a)\alpha_i}\right)^k \\ +\sum_{i=N_c}^{N} p_i^A \left(b - \sqrt{(b-a)(b-c)(1-\alpha_i)}\right)^k \end{matrix}, \quad \ldots\ldots\ldots\text{(A-20)}$$

where N_c is the approximation Point i at which $F(c) = \alpha_i$.

The raw moments are complex in the case of the triangular, and we restrict our attention to the case where $a = 0$. Writing c as a fraction z of b, the approximate kth raw moment is then

$$\hat{\mu}_k = b^k \begin{bmatrix} \sum_{i=1}^{N_c} p_i^A \left(\sqrt{z^{-1}\alpha_i}\right)^k \\ +\sum_{i=N_c}^{N} p_i^A \left(1 - \sqrt{(1-z^{-1})(1-\alpha_i)}\right)^k \end{bmatrix}. \quad \ldots\ldots\ldots\text{(A-21)}$$

The true kth raw moment is

$$\mu_k = \frac{2\left(bc^{k+2} - b^{k+2}c\right)}{b(b-c)c(2+3k+k^2)} = b^k \frac{2(z^{-k-2} - z^{-1})}{(z^{-1}-z^{-2})(k^2+3k+2)}. \quad \ldots\ldots\ldots\ldots\text{(A-22)}$$

The variance of the kth raw moment is

$$\sigma_k^2 = b^{2k}\frac{\begin{Bmatrix} k^2(5+k) + z^{-2k-1}\left[k^2(5+k) + 8(1+2k)\right] \\ -(z^{-1} + -z^{-2k-1})(1+k)(2+k)^2 \end{Bmatrix}}{z^{-2}(z-1)^2(1+2k)(2+3k+k^2)^2}. \quad \ldots\ldots\text{(A-23)}$$

Thus, we have

$$\delta_k = \frac{\mu_k - \hat{\mu}_k}{\sigma_k} = \frac{\begin{matrix} \frac{2(z^{-k-2} - z^{-1})}{(z^{-1}-z^{-2})(k^2+3k+2)} \\ -\left(\begin{matrix} \sum_{i=1}^{N_c} p_i^A \left(\sqrt{z^{-1}\alpha_i}\right)^k \\ +\sum_{i=N_c}^{N} p_i^A \left(1 - \sqrt{(1-z^{-1})(1-\alpha_i)}\right)^k \end{matrix}\right) \end{matrix}}{\sqrt{\frac{\begin{bmatrix} k^2(5+k) + z^{-2k-1}(k^2(5+k) + 8(1+2k)) \\ -(z^{-1} + -z^{-2k-1})(1+k)(2+k)^2 \end{bmatrix}}{z^{-2}(z-1)^2(1+2k)(2+3k+k^2)^2}}}, \quad \ldots\ldots\ldots\ldots\text{(A-24)}$$

and we see that δ is independent of b and that the *S*-equivalences that we list in Table 7 hold for any $T(0, b, b/2)$.

Log-Normal *S*-Equivalence. The inverse CDF for the log-normal is $F^{-1}(\alpha_i) = \exp[\mu + \sigma\Phi^{-1}(\alpha_i)]$. The approximate kth raw moment for approximation A is then

$$\begin{aligned} \hat{\mu}_k &= \sum_{i=1}^{N} p_i^A \left(\exp[\mu + \sigma\Phi^{-1}(\alpha_i^A)]\right)^k \\ &= \sum_{i=1}^{N} p_i^A \exp[k\mu + k\sigma\Phi^{-1}(\alpha_i^A)] \end{aligned}. \quad \ldots\ldots\ldots\ldots\text{(A-25)}$$

The true kth raw moment of the log-normal is

$$\mu_k = \exp\left[k\mu + \frac{1}{2}k^2\sigma^2\right]. \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(A-26)}$$

The variance of the kth raw moment is

$$\sigma_k^2 = \exp\left[2k\mu + k^2\sigma^2\right]\left(\exp\left[k^2\sigma^2\right] - 1\right). \quad \ldots\ldots\ldots\ldots\text{(A-27)}$$

Thus, we have

$$\delta_k = \frac{\mu_k - \hat{\mu}_k}{\sigma_k} = \frac{\exp\left[\frac{1}{2}k^2\sigma^2\right] - \sum_{i=1}^{N} p_i^A \exp[k\sigma\Phi^{-1}(\alpha_i^A)]}{\sqrt{\exp\left[k^2\sigma^2\right]\left(\exp\left[k^2\sigma^2\right] - 1\right)}}, \quad \ldots\ldots\ldots\ldots\text{(A-28)}$$

and we see that δ is independent of μ and that the *S*-equivalences that we list in Table 7 hold for any $L(\mu, 1)$.

J. Eric Bickel is an assistant professor in both the graduate program in operations research and the Department of Petroleum and Geosystems Engineering at the University of Texas at Austin. In addition, he is a fellow in the Center for Petroleum Asset Risk Management. Before returning to academia, Bickel was a Senior Engagement Manager for SDG. He holds a BS degree in mechanical engineering, with a minor in economics, from New Mexico State University and both MS and PhD degrees in the Engineering-Economics Systems program from Stanford University. Bickel can be reached at ebickel@mail.utexas.edu. **Larry W. Lake** is a professor in the Department of Petroleum and Geosystems Engineering at the University of Texas at Austin. He directs the Center for Petroleum Asset Risk Management. Lake holds BSE and PhD degrees in chemical engineering from Arizona State University and Rice University, respectively. **John Lehman** is a principal at SDG and director of the SDG-Galway Energy Practice. Previous to this, he was the chief risk officer of a power and natural-gas trading company. Lehman holds a BS degree in finance and economics from California State University Sacramento and an MBA degree, with emphases in finance and economics, from the University of Washington.