Decision Analysis
Vol. 10, No. 1, March 2013, pp. 6–25
ISSN 1545-8490 (print) | ISSN 1545-8504 (online)
http://dx.doi.org/10.1287/deca.1120.0260

Reexamining Discrete Approximations to Continuous Distributions

Robert K. Hammond, J. Eric Bickel
Graduate Program in Operations Research and Industrial Engineering, University of Texas at Austin, Austin, Texas 78712
{robertkh@mail.utexas.edu, ebickel@mail.utexas.edu}

Discretization is a common decision analysis technique for which many methods are described in the literature and employed in practice. The accuracy of these methods is typically judged by how well they match the mean, variance, and possibly higher moments of the underlying continuous probability distribution. Previous authors have analyzed the accuracy of differing discretization methods across a limited set of distributions drawn from particular families (e.g., the bell-shaped beta distributions). In this paper, we extend this area of research by (i) using the Pearson distribution system to consider a wide range of distribution shapes and (ii) including common, but previously unexplored, discretization methods. In addition, we propose new three-point discretizations tailored to specific distribution types that improve upon existing methods.

Key words: probability discretization; subjective probability; decision analysis; Pearson distribution; Swanson–Megill; practice
History: Received on July 2, 2012. Accepted by former Editor-in-Chief L. Robin Keller on November 9, 2012, after 1 revision. Published online in *Articles in Advance* January 18, 2013.

1. Introduction

In many decision analysis problems, uncertain quantities of interest are represented by continuous probability density functions (pdfs). To aid assessment, analysis, or explanation, these pdfs are often *discretized* into discrete probability mass functions (pmfs). Specifically, suppose we are constructing a decision model that takes as an input the continuous random variable X (e.g., oil reserves). We approximate the pdf $f(x)$, or $F(x)$; the cumulative distribution function (cdf) with a set of values $x_i \in X$, $i = 1, 2, \ldots, N$; and associated probabilities $p_i \equiv p(x_i)$. In most discretization methods, N is equal to three, but five is not uncommon. The values, which can be expressed as fractiles, or percentiles, of $F(x)$, and probabilities are selected such that desired properties of X are preserved. The most natural and common properties of interest are the raw and central moments (e.g., the mean, variance, skewness, and kurtosis).

Several discretization methods are in common use. For example, the extended Swanson–Megill (ESM) (Keefer and Bodily 1983) method weights the 10th (P10), 50th (P50), and 90th (P90) percentiles of $F(x)$ by 0.300, 0.400, and 0.300, respectively, and is heavily used within the oil and gas industry (Megill 1977, Hurst et al. 2000, Rose 2001, Bickel et al. 2011). Another common discretization is the "25–50–25" method, which weights the P10, P50, and P90 by 0.250, 0.500, and 0.250, respectively, and is also heavily used in the oil and gas industry. Companies tend to be committed to only one of these two methods, although many other methods exist. Keefer and Bodily (1983) (hereafter, KB) suggested treating the mean-approximation formula of Pearson and Tukey (1965) as a complete pmf and referred to this method as the extended Pearson–Tukey (EPT) method. EPT weights the P5, P50, and P95 by 0.185, 0.630, and 0.185, respectively. Additional discretizations and the history of their development will be discussed in §2.

The many different discretization methods naturally raise the question as to which method is best in general or for a particular situation. Miller and Rice (1983) showed that any discretization method whose values are interval means will underestimate the even moments of the underlying pdf. They proposed a powerful moment-matching discretization method based on Gaussian quadrature (GQ) that with n points, can match $2n$ moments of a pdf (assuming that those moments are finite), including the zeroth moment, which requires the probabilities to sum to one. Smith (1993) described this method in detail and compared it to several other methods. GQ should be considered the "gold standard" because it can match as many moments as desired. Yet, it is more difficult to apply than the methods discussed here. Please see Bickel et al. (2011) for GQ discretizations of common distributions, including the uniform, normal, exponential, triangular, and lognormal.

KB studied how well several discretization methods and approximation formulae matched the mean and variance of 78 bell-shaped beta distributions (labeled as ∩-shaped) and found EPT to be the best. Keefer (1994) examined the accuracy of approximating certain equivalents, which depend upon all the moments of the underlying pdf, and considered a wider set of 169 beta distributions (including the 78 considered by KB), covering symmetric and right- and left-skewed ∩-shapes, and both right- and left-skewed *J*-shapes, as well as four lognormal distributions. Keefer's (1994) analysis showed that EPT had the lowest error across several levels of risk aversion, including risk neutrality, in which case the certain equivalent is simply the mean.

Development of three-point approximation formulae for the mean and variance has also been pursued in the Program Evaluation Review Technique (PERT) literature. For example, Keefer and Verdini (1993) and Johnson (1998, 2002) studied the accuracy of existing three-point formulae, such as Pearson and Tukey (1965), and developed new formulae for particular distributions, respectively. Although this work is related to ours, it differs both in emphasis and scope. Our goal is the development of discrete approximations to continuous distributions, not formulae to approximate individual moments. In addition, rather than limiting our study to particular distributions such as the ∩-shaped beta or lognormal, we systematically consider a very wide range of distribution shapes through the use of the Pearson[1] (1895, 1901, 1916) distribution system, which was also used by Pearson and Tukey (1965).

[1] Karl Pearson (1857–1936) was the father of Egon Sharpe Pearson (1895–1980), of Pearson and Tukey (1965).

In this paper, we (1) extend previous research by considering a wider range of distributions, through the use of the Pearson distribution system; (2) analyze discretization methods that have not been previously considered, but that are now in common use; and (3) suggest several new discretizations that are tailored to specific distribution families. These discretizations are similar in spirit to EPT, but are optimized for a narrower set of distributions.

The remainder of this paper is organized as follows. The next section describes the discretization methods that we consider. Section 3 describes the Pearson distribution system, and serves as the foundation of our work. Section 4 presents our analysis of existing discretization methods. Section 5 presents several new discretization shortcut methods and analyzes their performance. Section 6 provides recommendations for practice and concludes.

2. Discretization Methods

We categorize discretization methods into *shortcuts*, which do not depend upon the shape of the underlying pdf, and *distribution specific*, which do. Shortcuts define a fixed set of percentiles to be used as discretization values, requiring only those percentiles to be known (or assessed). Distribution-specific methods, instead, may use different percentiles for different distributions, generally requiring that the full distribution be known. This section summarizes and discusses the development of the shortcuts and distribution-specific methods that we analyze in this paper.

2.1. Shortcuts

Pearson and Tukey (1965) tested many approximations that preserve either the mean or the variance of the underlying pdf, across a set of distributions drawn from the Pearson system. They settled upon a symmetric three-point approximation for the mean and a system of equations to approximate the variance. Later, KB suggested treating Pearson and Tukey's (1965) mean approximation as a full pmf, which is more useful than a mean approximation in a decision analysis context, and referred to it as the extended Pearson–Tukey. As described in §1, EPT weights the P5, P50, and P95 by 0.185, 0.630, and 0.185, respectively. As a summary, we will write discretizations of this form as (P5, P50, P95, 0.185, 0.630, 0.185).

Roy Swanson proposed, in a 1972 internal Exxon memo, that the mean of a lognormal distribution can be approximated by weighting the P10, P50, and P90 by 0.300, 0.400, and 0.300, respectively (Hurst et al. 2000). Megill (1977) was the first to publish "Swanson's mean" and stressed that it should not be used for highly skewed lognormal distributions. KB suggested treating Swanson and Megill's mean approximation as a full pmf and called it *extended Swanson–Megill*, to be used with many distribution families, not just the lognormal. Recently, Bickel et al. (2011) showed that ESM can be inaccurate for even moderately skewed distributions and instead recommended the use of GQ.

Miller and Rice (1983) introduced the use of GQ (described in §2.2) to determine discretizations that perfectly match the moments of the underlying pdf. This method works well when the pdf is known to be from a specified family. To address the case where the underlying pdf is not from a known family, as might be the case when one assesses a cdf directly from an expert, Miller and Rice (1983) proposed the (P8.5, P50, P91.5, 0.248, 0.504, 0.248) shortcut, known as the Miller and Rice one-step (MRO). Given the similarity of the MRO to the McNamee–Celona shortcut (MCS), which we describe shortly, we do not analyze the MRO in this paper. However, a MRO analysis is provided in the online supplement to this paper (available at http://dx.doi.org/10.1287/deca.1120.0260).

D'Errico and Zaino (1988) proposed, and Zaino and D'Errico (1989) analyzed, two approximations based on Taguchi's (1978) experimental-design method. The first (Zaino–D'Errico "Taguchi" (ZDT)) uses equal weights (P11, P50, P89, 0.333, 0.333, 0.333). The second (Zaino–D'Errico "Improved" (ZDI)) is a three-point Gaussian quadrature for a normal distribution, which is (P4.2, P50, P95.8, 0.167, 0.667, 0.167). Zaino and D'Errico (1989) found that ZDI was more accurate, in many situations, than ZDT, and we therefore do not consider ZDT in this paper. However, an analysis of ZDT is included in the online supplement.

McNamee and Celona (1990) described another shortcut that has come to be known as the McNamee–Celona shortcut, or the "25–50–25" shortcut. The MCS uses (P10, P50, P90, 0.250, 0.500, 0.250). It is based on both the MRO and the application of a distribution-specific method known as bracket mean (BMn).

Table 1 **Summary of Shortcut Discretization Methods Considered in This Paper**

Shortcut	Percentile points	Probability weights
Extended Pearson–Tukey	P5, P50, P95	0.185, 0.630, 0.185
Extended Swanson–Megill	P10, P50, P90	0.300, 0.400, 0.300
McNamee–Celona shortcut	P10, P50, P90	0.250, 0.500, 0.250
Zaino–D'Errico "Improved"	P4.2, P50, P95.8	0.167, 0.667, 0.167

McNamee and Celona (1990) cautioned that the MCS was only a first approximation in analyzing a decision problem and that the distributions should be encoded and discretized more carefully as the analysis progressed. This recommendation is not always followed in practice (Bickel et al. 2011).

Table 1 summarizes the discretization shortcuts considered in this paper. KB did not consider the MCS or ZDI. Keefer (1994) considered all four of these shortcuts (as well as the MRO and ZDT) and also investigated a wider range of distributions than KB, comprised of 169 beta and four lognormal distributions. However, as will be explained in §3, Keefer's analysis range is much narrower than what we consider here.

2.2. Distribution-Specific Methods

Distribution-specific methods require specifying the pdf (or the cdf) to be discretized, instead of only three percentiles. We consider two discretization-specific methods: bracket mean, also known as equal areas (Bickel et al. 2011), and bracket median (BMd). BMn was originally developed by Jim Matheson and his colleagues (Bickel et al. 2011) at the Stanford Research Institute between the late 1960s and the early 1970s. Both methods horizontally divide the cdf into intervals (three is common, but not necessary). These intervals could be equal but seldom are in the case of BMn. For example, the most common method is to divide the cdf into intervals between the P100 and the P75, the P75 and the P25, and the P25 and the P0. This produces a weighting of 0.25, 0.50, and 0.25, respectively. BMd summarizes each interval with the conditional median of that interval, whereas BMn uses the conditional mean. Because these conditional distributions are generally skewed, the median and the mean differ and the two approaches may result in different discretizations. Applying the three-point BMn method with intervals of 0.25, 0.50, and 0.25 to the normal distribution yields the discretization (P12.5, P50.0, P87.5, 0.25, 0.50, 0.25). That this BMn discretization has the same probabilities and similar percentiles as the MCS lends additional support to the MCS.

We consider three- and five-point versions of bracket median (BMd3 and BMd5) and bracket mean (BMn3 and BMn5). In the case of BMd3, BMd5, and BMn5, we divide the cdf into equal probability intervals. In the case of BMn3, we follow standard practice (McNamee and Celona 1991) and choose 25–50–25 intervals. Of these four methods, KB considered only BMd5.

3. The Pearson Distribution System

Several distribution systems that collectively define sets of continuous distributions have been developed. The best known is the Pearson system, which was described by Karl Pearson in a series of papers (Pearson 1895, 1901, 1916).[2] A distribution $f(x)$ in the Pearson system is a solution of the differential equation

$$\frac{1}{f}\frac{df}{dx} = \frac{b-x}{c_0 + c_1 x + c_2 x^2}. \tag{1}$$

The four parameters in Equation (1), b, c_0, c_1, and c_2, determine the first four moments of the associated pdf. Together, within the Pearson system, the third and fourth moments determine a unique location-scale distribution (see, for example, Elderton and Johnson 1969).

This latter fact allows distributions in the Pearson system to be conveniently characterized by their shape, which is defined by their skewness γ_1 and kurtosis β_2 (the third and fourth central moments, respectively). Because skewness can be positive or negative, but is symmetric under reflection, it is convenient to consider squared skewness, $\beta_1 = \gamma_1^2$.

[2] Other distributions systems include those by Johnson (1949), Burr (1973), Ramberg and Schmeiser (1974), Schmeiser and Deutsch (1977), and Butterworth (1987). We do not use these systems for a variety of reasons. Johnson's (1949) system does not reduce to as many common distributions as Pearson's. Burr's (1973) system and Schmeiser and Deutsch's (1977) system each have only a single range of support, versus the three of Pearson's system. Ramberg and Schmeiser's (1974) system covers a smaller range of shapes than Pearson's. Butterworth's (1987) system only approximates several of the named distributions included in the Pearson system.

Figure 1 Pearson Distribution System

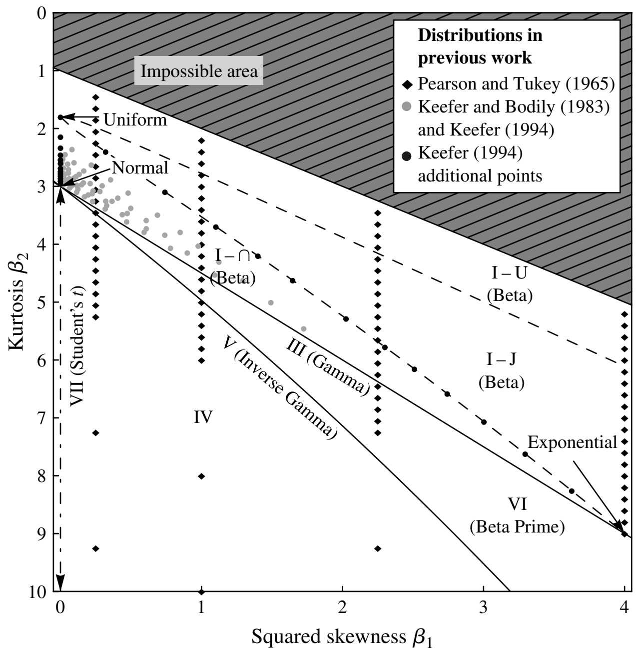


Figure 1 shows a portion of the Pearson system, denoting several regions, or classes of distributions. The β_1 and β_2 axes are not bounded above. The vertical axis denoting kurtosis is inverted following the tradition of previous work.

The Pearson system includes all possible combinations of skewness and kurtosis, which Pearson (1916) showed must have

$$\beta_2 \geq \beta_1 + 1. \tag{2}$$

The region above the line $\beta_2 = \beta_1 + 1$ is shaded and labeled as the "Impossible area" in Figure 1. Although the Pearson system covers all possible (β_1, β_2) pairs, it does not include all possible pdfs, most notably the lognormal (included in the Johnson 1949 system). However, as highlighted in Figure 1, many common distributions are special cases, including the beta, uniform, normal, exponential, gamma, and inverse gamma.

Lau et al. (1998) (hereafter, LLH) used a combination of the Pearson system and three others (Johnson 1949, Ramberg and Schmeiser 1974, Schmeiser and Deutsch 1977) to design and evaluate several mean-approximation formulae over the parameter range $\beta_1 < 4$, $\beta_2 < 9$. We use a similar but slightly larger range, $\beta_1 < 4$, $\beta_2 < 10$, to increase coverage of kurtosis and include more of the distributions considered by Pearson and Tukey (1965). This range is arbitrary, but through experimentation we determined that increasing this range by 50% in each dimension does not change our conclusions. The Pearson system is a convenient and natural choice because of its direct relation to named distributions over much of the feasible (β_1, β_2) region, its flexibility in distribution shape and support, and general ease of use as compared to other systems.

Three main types of distributions cover the feasible region, which Pearson designated Type I, Type IV, and Type VI. Pearson defined nine additional types, which are special cases of the main three or transition boundaries between them. For example, Types III and V are the gamma and inverse gamma distributions, respectively. The normal distribution is a special case, where Types I, II (not shown), III, IV, V, and VI intersect.

The distributions in Figure 1 that were considered by both KB and Keefer (1994) are denoted by grey circles, as shown in the figure's legend. Distributions considered by Keefer (1994), but not by KB, are denoted by black circles. Both KB and Keefer (1994) considered a relatively small sample of distributions within the Pearson system, being confined to ∩-shaped beta (Type I) distributions with low skew. Keefer (1994) expanded this somewhat, but limited the analysis to beta distributions along the transition between the ∩-shaped and J-shaped betas regions and four lognormal distributions. The possible (β_1, β_2) points for the lognormal fall within the Pearson Type VI region, and although the lognormal distribution is not included in the Pearson system, our analysis of this region covers distributions of similar shape.

The distributions Pearson and Tukey (1965) used to construct their mean- and variance-matching formulae are shown as black diamonds in Figure 1. Although 11 of the 96 points they used fall outside the area of Figure 1, these points are sparsely distributed over a range of kurtoses from 10 to 20 and are inconsistent with the denser spacing of the rest of their grid. Pearson and Tukey's (1965) analysis did not fully explore the Pearson system, because it was limited to the tables of Pearson distributions that were available at the time.

We now briefly describe the three main Pearson distributions and two transition distributions in the top-to-bottom order in which they appear in Figure 1. The distributions in the Pearson system are location-scale generalizations, but we give the standard forms of the distributions, which are equivalent under an appropriate shifting and scaling of x.

3.1. Pearson Type I (Beta Distribution)

Type I corresponds to the beta distribution, with pdf

$$f_I(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1, \tag{3}$$

where α and β are parameters and $B(\alpha, \beta)$ is the beta function. Pearson (1895) characterized this type as having skewness and limited range. This type arises when the denominator of Equation (1) has roots with opposite signs. The beta can be ∩-shaped ($\alpha \geq 1$, $\beta \geq 1$), J-shaped ($\alpha \geq 1, \beta < 1$ or $\alpha < 1$, $\beta \geq 1$), or U-shaped ($\alpha < 1, \beta < 1$). We denote these shapes as Types I-∩, I-J, and I-U, respectively. When a beta distribution is J-shaped, the value of f approaches infinity as x approaches 0 (when $\alpha \geq 1, \beta < 1$) or 1 (when $\alpha < 1, \beta \geq 1$). When it is U-shaped, f goes to infinity at both 0 and 1. Type I-U is the only Pearson type that is not unimodal. The symmetric Type I is called Type II and lies along the β_2 axis (not shown) between kurtoses 1 and 3. The uniform distribution is the point $(\beta_1 = 0, \beta_2 = 1.8)$, which is also the point where the three regions of Type I meet.

3.2. Pearson Type III (Gamma Distribution)

Type III, or the gamma distribution, is a transition distribution that forms the boundary between Type I and Type VI in Figure 1. It has pdf

$$f_{III}(x) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-x/\theta}, \quad 0 \leq x < \infty, \tag{4}$$

where $\Gamma(k)$ is the gamma function, and k and θ are parameters. This type occurs when $b_2 = 0$ in Equation (1). At the point where this type intersects with the line that divides the ∩-shape and J-shape Type I regions, is the exponential distribution, or Type X.

3.3. Pearson Type VI (Beta Prime Distribution)

Type VI corresponds to the beta prime distribution, also called the inverted beta distribution or the beta distribution of the second kind. For parameters α and β, the pdf is

$$f_{VI}(x) = \frac{x^{\alpha-1}(1+x)^{-\alpha-\beta}}{B(\alpha, \beta)}, \quad 0 < x < \infty. \qquad (5)$$

Pearson (1901) characterized this distribution as being unbounded in one direction. As seen in Figure 1, Type VI covers the region between the gamma and inverse gamma distributions, each of which also have this property. Type VI distributions are the solution to Equation (1) when its denominator has roots of the same sign.

3.4. Pearson Type V (Inverse Gamma Distribution)

Type V is the second transition type that separates the regions of Type IV and Type VI and is known as the inverse gamma distribution. For parameters α and β, it has pdf

$$f_V(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}, \quad 0 < x < \infty. \qquad (6)$$

This type occurs when the roots of Equation (1) are real and equal.

3.5. Pearson Type IV

Type IV does not correspond to any single common distribution. For parameters m and v, the pdf is

$$f_{IV}(x) = y_0(1+x^2)^{-m} e^{-v \tan^{-1}(x)}, \quad -\infty < x < \infty, \qquad (7)$$

where y_0 is a normalizing factor (see Elderton and Johnson 1969). Pearson (1895) characterized this type as being unbounded in both directions and possibly having skewness. It is the solution to Equation (1) that arises when the denominator has complex roots. A special case when Type IV is symmetric ($\beta_1 = 0$, $\beta_2 > 3$) is Student's t-distribution, which Pearson (1916) called Type VII.

4. Discretization Accuracy

As discussed in §3 and shown in Figure 1, KB considered a small portion of the Type I-∩ (∩-shaped beta) distributions when analyzing the accuracy of three-point approximations, including EPT and ESM. This section extends their analysis by using the Pearson system to consider a wider range of distribution shapes and support types. In addition, we consider the MCS, ZDI, BMd, and BMn methods.

4.1. Methodology

We construct a grid of approximately 2,800 points covering the feasible region shown in Figure 1, spaced 0.1 in each dimension. LLH used 1,000 uniformly distributed points over a slightly smaller region, whereas Pearson and Tukey (1965) used a grid of 96 points, as described in §3 and shown in Figure 1. Without loss of generality, we consider only positively skewed, or right-skewed, distributions. With the mean and variance normalization we discuss next, the direction of skew does not affect the error.

For each distribution represented by a point in our grid, we construct the discretization and measure the error between the mean and variance of the discretization and the actual mean and variance, respectively, of the associated distribution. The shortcut methods and BMd require only that points be taken from the cdf, whereas BMn requires the computation of conditional means, and entails numerically integrating regions of the pdfs given in §3. Numerical integration introduces numerical error, but this error will turn out to be negligible.

4.2. Analysis of the Beta Distribution

To illustrate our technique, we begin by benchmarking it against the results of KB, who focused on Type I-∩. KB measured performance by considering four error measures: average absolute error (AAE), average absolute percentage error (AAPE), maximum error (ME), and maximum percentage error (MPE). ME and MPE give the error with the greatest magnitude, regardless of sign. For these two measures, positive error indicates that the discretized moment is larger than the actual moment. We define these error measures as follows: Let D_r denote the indexed set of distributions corresponding to region $r \in \{\text{I-}\cap, \text{I-J}, \text{I-U}, \text{III}, \text{IV}, \text{V}, \text{VI}\}$. The number of points that we sample in this region is s_r. For a distribution index $i \in D_r$, the true kth moment is m_i^k, the estimate from a discretization is $\hat{m}_i^k$, and the *error* in the kth moment is $\varepsilon_i^k = (\hat{m}_i^k - m_i^k)$. Table 2 provides the formulae for error measures considered by KB.

In Table 3, we report the results for the I-∩ distributions, in the same manner as KB. A selection of KB's results is repeated in Table 4 for comparison. Although each of EPT, ZDI, BMn3, and BMn5 has at least one measure of "0.000" for the mean in Table 3,

Table 2 KB Error Measures

Error statistic	Formula
Average absolute error	$\frac{1}{s_r}\sum_{i\in D_r}\lvert\varepsilon_i^k\rvert$
Average absolute percentage error	$\frac{1}{s_r}\sum_{i\in D_r}\left\lvert\frac{\varepsilon_i^k}{m_i^k}\right\rvert\cdot 100\%$
Maximum error	$\varepsilon_j^k,\ j=\arg\max_{i\in D_r}\lvert\varepsilon_i^k\rvert$
Maximum percentage error	$\frac{\varepsilon_j^k}{m_j^k}\cdot 100\%,\ j=\arg\max_{i\in D_r}\left\lvert\frac{\varepsilon_i^k}{m_i^k}\right\rvert$

they are not exactly equal to zero. The highlighted cells show the most accurate shortcut (top rows of Table 3) and the most accurate distribution-specific method (bottom rows of Table 3) for each measure. As noted by KB, EPT performs very well. ZDI has comparable performance to EPT and is better than ESM. BMd performs rather poorly. The BMn3 and BMn5 methods perfectly match the mean (with negligible numerical error), but are not as accurate as EPT in matching the variance. The errors for the MRO and ZDT are given in the online supplement.

All of the error measures are increased relative to KB's results because of our expanded distribution set. For example, EPT's AAPEs in the mean and the variance are about three times larger in our case than in the case reported by KB (0.066% compared to 0.020%). In addition, our expanded analysis demonstrates that EPT outperforms ESM more than was found by KB. For example, in our case, ESM's AAPE is about five times as large as EPT's, whereas KB found that it was only about two and one-half times as large.

BMn, in theory, matches the means of all of the distributions considered, but in practice the integrals involved in computing the conditional means often can be evaluated only using numerical integration methods, which introduces numerical error. However, good numerical integration software will generally produce negligible errors,[3] which, in our case, are several orders of magnitude smaller than the best discretization errors. Also, in practical application, GQ will often have small numerical errors.

[3] We used numerical integration algorithms included in the Mathematica 8 software package.

4.3. Expanded Distribution Set Using Pearson's System

We turn now to the expanded set of distributions from the Pearson system. The beta distribution has the property that $F(\alpha,\beta)=1-F(\beta,\alpha)$. However, percentage error does not follow this relation for nonstandardized beta distributions and introduces bias, as pointed out by LLH. LLH standardized each distribution to have unit mean and unit standard deviation. This practice, which we follow, eliminates the bias and allows for consistent comparison of errors between distributions with different support ranges (e.g., the Beta (Type I) and Beta Prime (Type VI)).

4.3.1. Error of Discretization Shortcuts. For each of the discretization shortcuts considered in this paper, we plot the absolute errors in the mean $\lvert\varepsilon_i^1\rvert$ and variance $\lvert\varepsilon_i^2\rvert$ over the entire region in Figure 1. Error in the mean is shown in Figure 2, and error in the variance in Figure 3. The MRO and ZDT error plots are included in the online supplement.

These plots quantify the performance of the various methods as a function of distribution shape. The contours and error magnitude ranges are standardized separately for the mean and variance throughout the error plots. Absolute error in the mean is shown to vary from 0 to 0.15, and in the variance from 0 to 0.5 (with our $\mu=\sigma=1$ normalization, these are equivalent to 0%–15% and 0%–50% error ranges, respectively), with darker shading indicating higher absolute error. Black areas in the plots indicate where the errors are greater than the upper bound.

With the exception of Type I-U, and I-J in the case of ZDI, EPT and ZDI perform well over most of the plot area. ESM displays greater errors than EPT and ZDI over Types I-U and I-J, and portions of the Type IV, V, and VI distributions. The MCS results in the highest error levels among the four shortcut methods. It performs well for the normal distribution, on which it is based, but its error increases rapidly with skewness. Both ESM and the MCS display significantly higher error sensitivity to distribution shape than either EPT or ZDI, and error generally increases with kurtosis and skewness. However, the MCS is much more sensitive to skewness than ESM because the MCS places less weight on the P10 and P90. ESM clearly outperforms the MCS for Type I-∩, III, IV, V, and VI distributions. All of the shortcut methods produce large errors

Table 3 **Results for Expanded Set of I-∩ (Beta) Distributions (Discretized Minus Actual Mean and Variance)**

	Mean				Variance			
	AAE	AAPE	ME	MPE	AAE	AAPE	ME	MPE
I-∩ (Beta)								
Shortcuts methods								
EPT	0.000	0.066	−0.001	−0.178	0.000	1.096	−0.008	−10.090
ZDI	0.000	0.152	−0.002	−0.589	0.000	1.189	−0.013	−16.094
ESM	0.001	0.331	0.004	1.419	0.001	7.776	0.013	−19.960
MCS	0.002	2.255	−0.005	−4.852	0.002	20.384	−0.005	−32.051
Distribution-specific methods								
BMd3	0.005	5.124	−0.012	−10.494	0.004	40.390	−0.012	−53.752
BMd5	0.003	3.027	−0.006	−6.390	0.002	25.961	−0.007	−38.402
BMn3	0.000	0.000	0.000	0.000	0.002	22.201	−0.013	−29.199
BMn5	0.000	0.000	0.000	0.000	0.001	13.042	−0.004	−20.526

Note. Shaded areas are the most accurate shortcuts (top) or distribution-specific methods (bottom).

Table 4 **Selected Results from KB (Discretized Minus Actual Mean and Variance)**

	Mean				Variance			
	AAE	AAPE	ME	MPE	AAE	AAPE	ME	MPE
I-∩ (Beta)								
EPT	0.000	0.020	0.000	0.070	0.000	0.460	−0.001	−1.600
ESM	0.000	0.050	0.001	0.330	0.000	2.700	0.006	11.100
BMd5	0.001	0.750	−0.004	−3.350	0.002	21.500	−0.006	−30.200

Note. Shaded areas are the most accurate methods.

within the I-U region, which is a strong indication that they should not be used for this Pearson type.

As shown in Figure 3, error in the variance increases distinctly with kurtosis, i.e., as the tails of the distributions get "fatter." EPT and ZDI are more sensitive to changes in skewness than either ESM or the MCS. However, with the exception of Type I-U, EPT and ZDI generally match the variance better than either ESM or the MCS. This undoubtedly stems from the fact that both EPT and ZDI's percentiles capture more of the tail effects than do ESM's and the MCS's P10 and P90 percentiles.

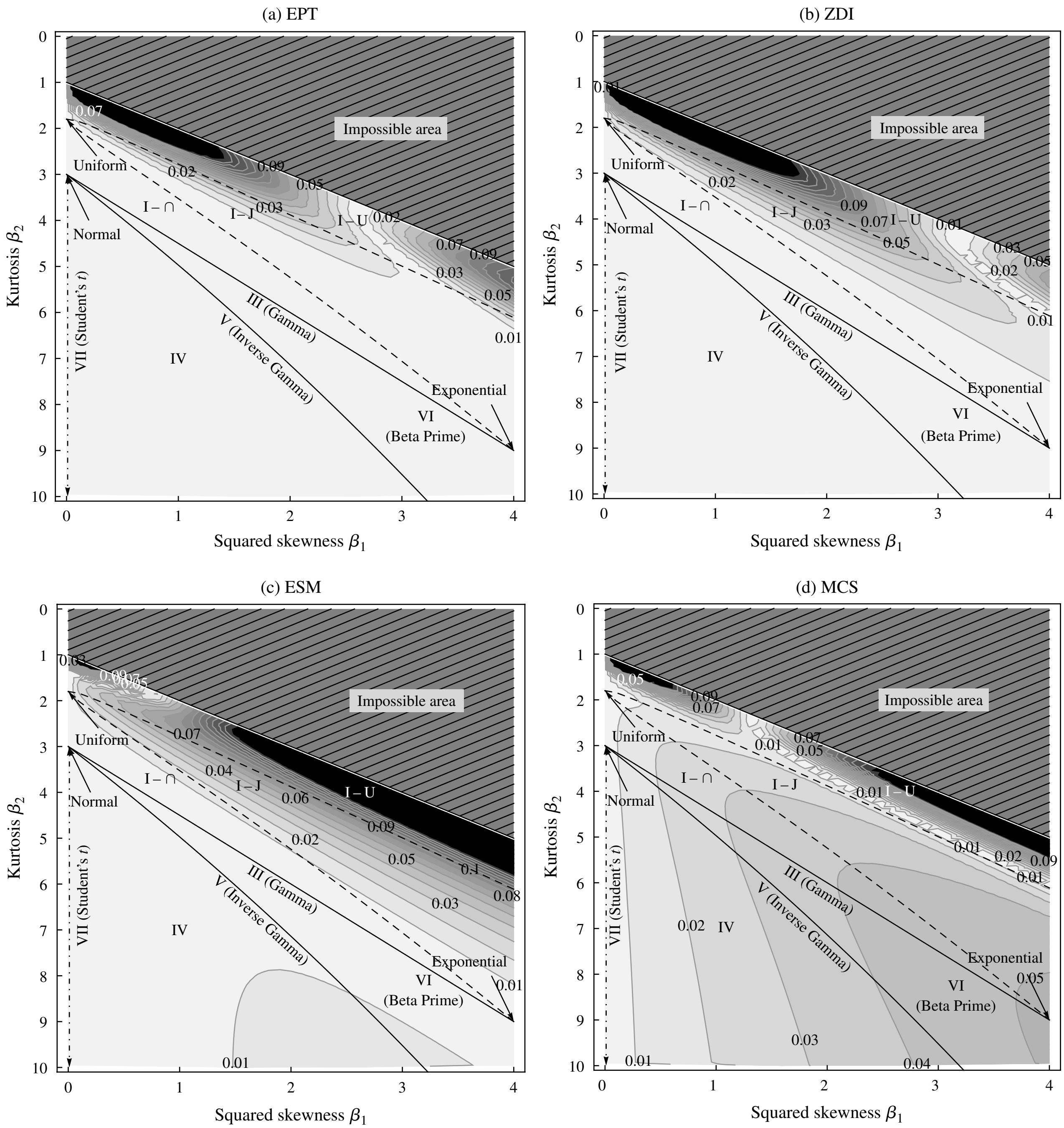
EPT and ZDI perform similarly over the Type I region, but ZDI more accurately matches the variance for Type IV, V, and VI distributions. The similarity in the EPT and ZDI errors is expected, considering the similarity of these methods' percentiles and probabilities. ZDI performs better over the Type IV distributions in Figure 3, perhaps because it is a GQ for the normal distribution, which is unbounded, as are the distributions in this region. Additionally, the formula used to derive EPT was designed by Pearson and Tukey (1965) based on performance over shapes mostly located in the Type I and VI regions (see Figure 1).

Although ESM's performance is rather poor, it is superior to that of the MCS. For example, although ESM misestimates the variance of Type I-∩ in the middle of our plots ($\beta_1 = 2$ and $\beta_2 = 5$) by about 10%, the MCS's error rate is 25%.

Before summarizing these results by Pearson type, as in Table 3, we introduce two new error measures: the average error (AE) and the average squared error (ASE), given in Table 5. The errors ε_i^k and percentage errors ε_i^k/m_i^k are the same for the standardized distributions.

Table 6 summarizes our results using only the AE, ASE, and ME for each of the main Pearson types (I, VI, and IV) over the skewness–kurtosis range depicted in our Pearson plots. Again, the difference between these types is their support range. As before, the best measures are highlighted for each distribution type. In Table 6, ASE's results are shown in scientific notation rather than rounded to three decimal places, because many would round to 0.000. These results identify the method that performs

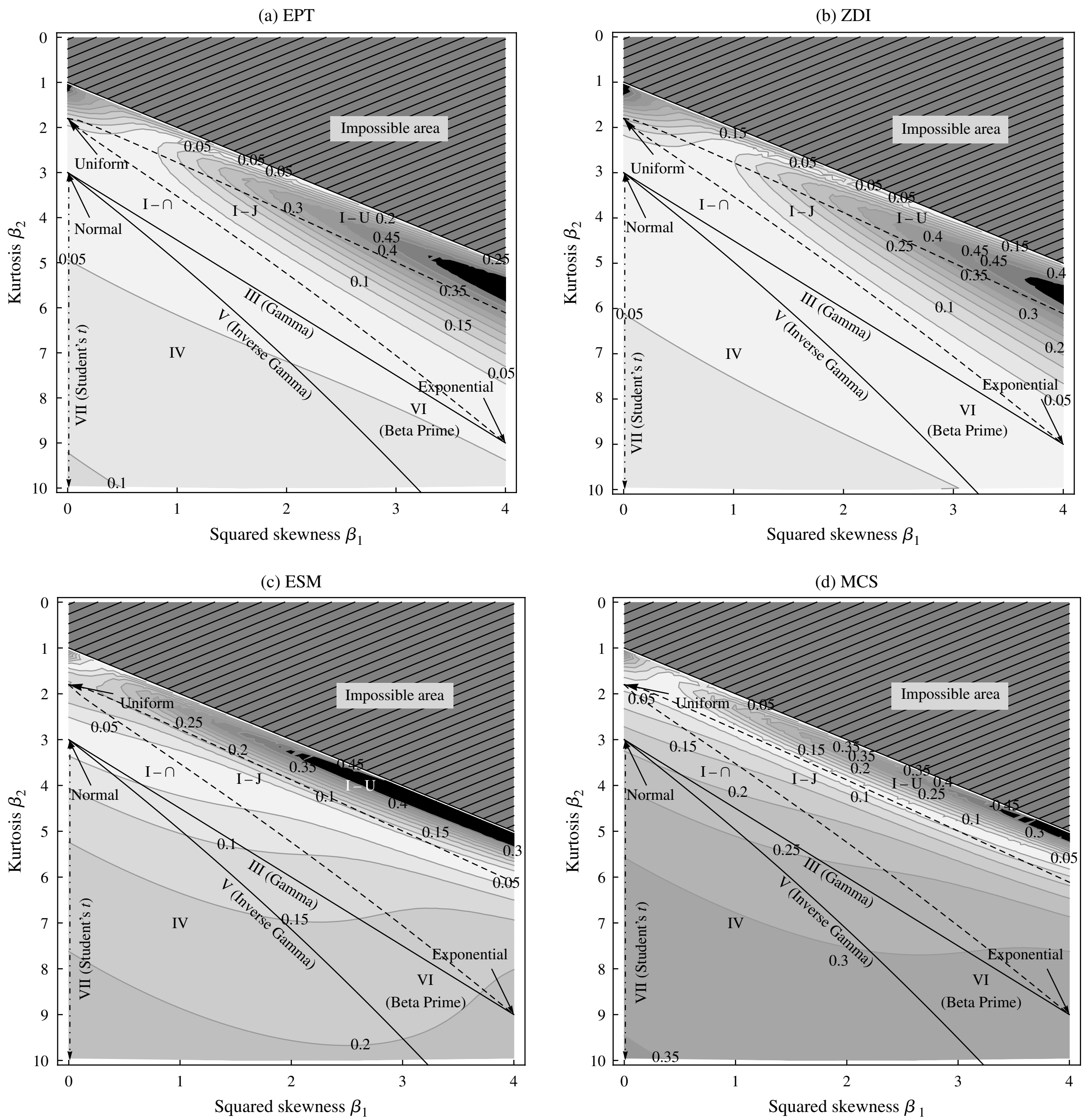
Figure 2 Errors in the Mean for Each Discretization Shortcut Method



best, when performance is averaged over a given distribution region. Errors and discussion for transition Types III and V, and the MRO and ZDT shortcuts, are included in the online supplement.

Either EPT or ZDI performs best for Type I-∩, IV, and VI distributions. These conclusions are consistent with the error plots of Figures 2 and 3. The I-U distribution is the only type for which the MCS

Figure 3 **Errors in the Variance for Each Discretization Shortcut Method**



performs best by any measure. ESM outperforms EPT only for the Type I-J variance and is the only case where ESM is best. All four shortcuts tend to underestimate the mean and variance for Types VI and IV, and both have at least one unbounded tail.

These results indicate that the method that is best at matching the mean is not necessarily the best at matching the variance. This is particularly apparent with EPT for Types I-J and VI. However, in these cases, EPT's variance-matching accuracy is nearly as good

Table 5 Additional Error Measures

Error statistic	Formula
Average error	$\frac{1}{s_r}\sum_{i\in D_r}(\varepsilon_i^k)$
Average squared error	$\frac{1}{s_r}\sum_{i\in D_r}(\varepsilon_i^k)^2$

as the accuracy of the shortcut that best matches the variance. Although EPT was designed to approximate only the mean, it preserves the variance better than, or nearly as well as, any of the shortcut methods.

Based on this analysis, we conclude that EPT is generally the best shortcut method for matching the mean and variance. ZDI's performance improvement over EPT for beta prime and Type IV is seldom significant. None of these shortcuts should be used for Type I-U distributions. Next, we analyze the distribution-specific methods.

4.3.2. Error of Distribution-Specific Discretizations. Because the BMn methods exactly match the mean, Figure 4 only presents BMd's error in estimating the mean. The contour levels are the same as those in Figure 2. BMd3 and BMd5 produce significant errors over most of the region we consider. This occurs because the conditional distributions are skewed, and therefore the conditional median is not equal to the conditional mean. Below the Type I-J region, error increases with skewness. Adding more points clearly improves the performance of BMd, although it is still inferior to the shortcut methods. For example, BMd5's performance resembles that of the MCS, but it is still slightly worse.

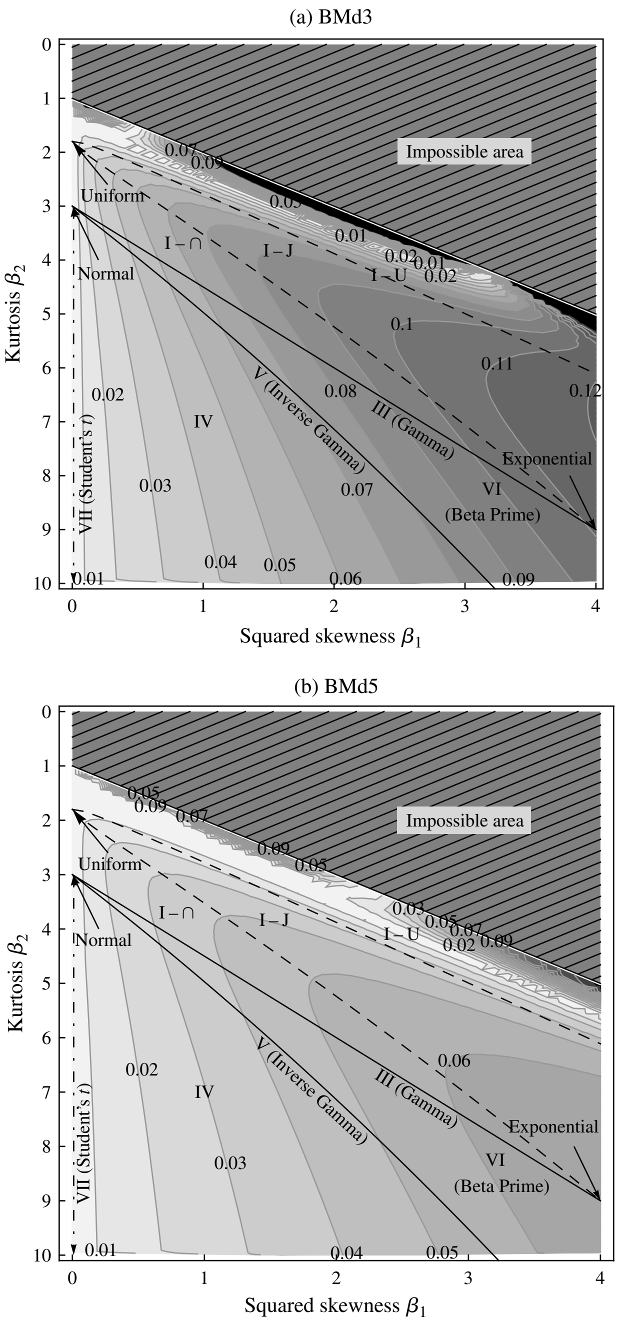
Figure 5 shows the variance-error plots for each of the BMd and BMn methods. The contour levels are the same as those in Figure 3. Error in the variance is primarily a function of kurtosis. BMd's performance is especially poor, although error is reduced by adding more discretization points. Yet, even a five-point BMn can underestimate the variance of low-skew Type I-∩ distributions by more than 10%. This suggests that a different discretization approach may be needed if preserving the variance is important. Even though BMd and BMn are tailored to the

Table 6 Shortcut Method Errors for Main Pearson Types (Discretized Minus Actual Mean and Variance)

	Mean			Variance		
	AE	ASE	ME	AE	ASE	ME
I-U (Beta)						
EPT	−0.041	1.15E−02	−0.412	0.214	1.05E−01	0.579
ZDI	−0.082	1.63E−02	−0.450	0.154	8.84E−02	−0.614
ESM	0.153	3.64E−02	0.449	0.304	1.29E−01	0.800
MCS	0.040	1.11E−02	0.302	0.155	5.61E−02	0.607
I-J (Beta)						
EPT	−0.004	9.46E−05	−0.044	0.100	2.02E−02	0.405
ZDI	−0.017	4.87E−04	−0.071	0.120	2.58E−02	0.441
ESM	0.034	1.68E−03	0.096	−0.067	1.39E−02	−0.212
MCS	−0.035	1.36E−03	−0.051	−0.197	4.58E−02	−0.330
I-∩ (Beta)						
EPT	0.001	9.75E−07	−0.003	−0.003	2.23E−04	−0.101
ZDI	−0.002	6.05E−06	−0.008	0.000	4.25E−04	−0.161
ESM	0.004	4.06E−05	0.019	−0.053	8.73E−03	−0.200
MCS	−0.030	1.05E−03	−0.050	−0.204	4.53E−02	−0.321
VI (Beta Prime)						
EPT	−0.001	8.92E−07	−0.002	−0.045	2.32E−03	−0.076
ZDI	−0.001	2.03E−06	−0.002	−0.021	5.89E−04	−0.048
ESM	−0.006	4.46E−05	−0.012	−0.166	2.93E−02	−0.221
MCS	−0.042	1.84E−03	−0.051	−0.298	9.01E−02	−0.341
IV						
EPT	−0.002	1.62E−05	0.003	−0.068	5.18E−03	−0.104
ZDI	−0.001	1.61E−06	0.007	−0.047	2.51E−03	−0.078
ESM	−0.007	6.01E−05	0.012	−0.173	3.12E−02	−0.224
MCS	−0.022	6.09E−04	0.044	−0.309	9.66E−02	−0.353

Note. Shaded areas are the most accurate for specific Pearson types.

Figure 4 Errors in the Mean for Bracket Median Methods



underlying distribution, the errors in the variance for each of these methods, over most of the region we consider, are significantly larger than those for EPT and ZDI. However, the BMn methods exhibit more gradual increases in error in and around the Type I-U region than do the shortcut methods. These results, averaged over our regions of interest, are summarized by distribution type in Table 7. Again, the errors for Types III and V for the distribution-specific methods are included in the online supplement.

As expected, the BMn methods are superior to both BMd and the shortcuts in matching the mean. Table 7 indicates that all of the numerical integration errors for the BMn methods average less than 10^{-3} and are generally several orders of magnitude smaller than the discretization errors for the BMd and shortcut methods. However, BMn systematically underestimates the variance, a general result that Miller and Rice (1983) proved. Here we show that the errors can be quite significant and are dependent on shape.

Except for Type I-U distributions, BMn results in larger errors in the variance than do the best shortcuts shown in Table 6. However, Table 7 and Figure 5 imply that the BMn errors are far less sensitive to distribution shape than are those for the shortcut methods, which, as seen in Figure 3, vary widely depending on distribution shape and type.

5. Best Mean Approximations by Pearson Type

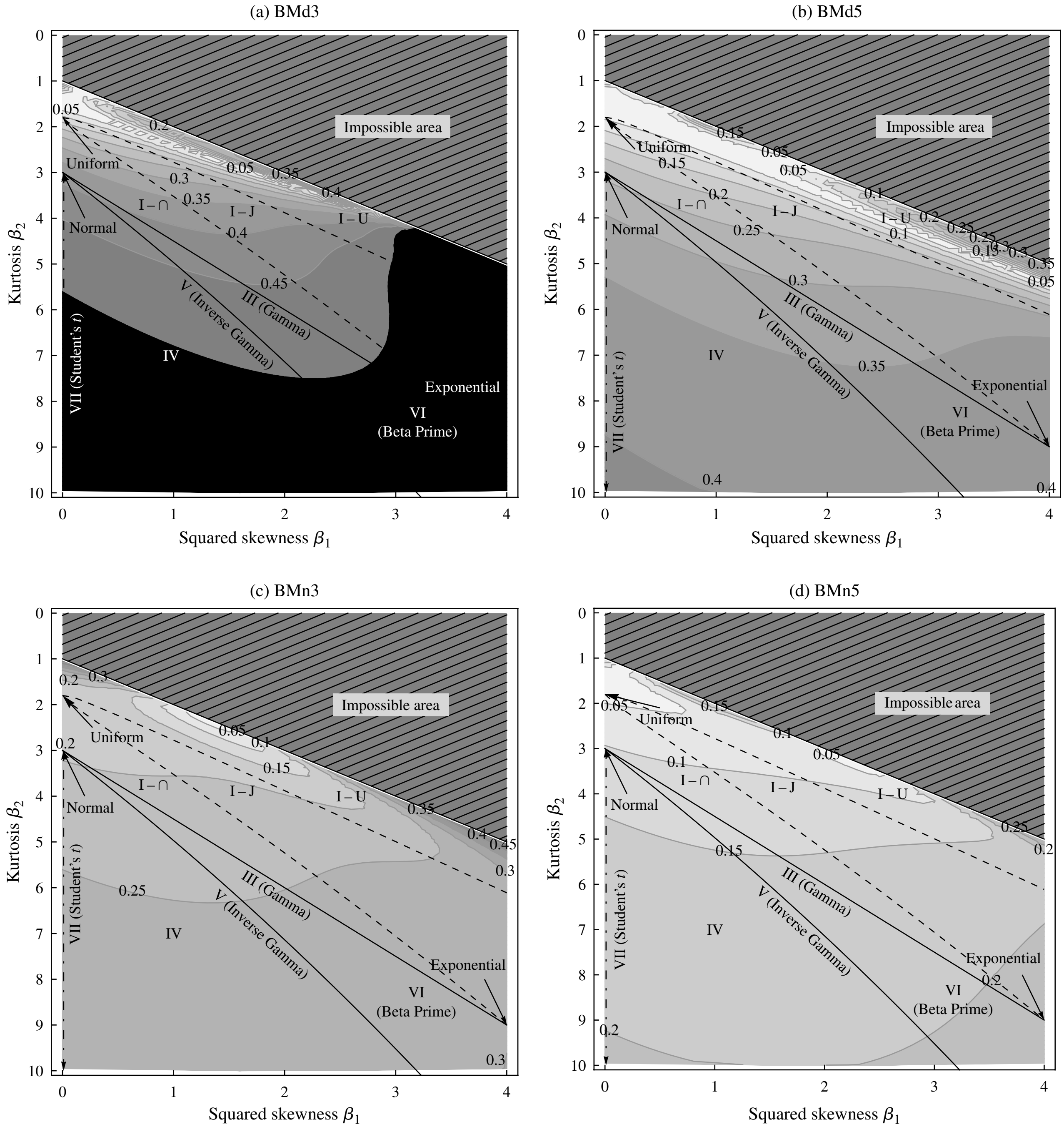
LLH found three-, five-, and seven-point mean-approximation formulae by finding, for a given set of percentiles, the probabilities that minimize the average squared error over the entire set of distributions (i.e., all the Pearson types). Here, we use a similar approach to find three-point mean approximation formulae for *individual* regions of the Pearson system. The primary distinction among the three major regions (i.e., I, VI, and IV) of the Pearson system is their range of support: Type I is bounded in both directions, Type VI is unbounded in one direction, and Type IV is unbounded in both directions. This basic characteristic of uncertainty is often easy to determine, allowing one to focus on a particular Pearson region and thereby choose the best shortcut.

LLH found their three-point approximations by solving

$$\min_{p_1,p_2} \ \frac{1}{n}\sum_{i=1}^{n}(\hat{m}_i^1 - m_i^1)^2$$

$$\text{s.t.}\ \ \hat{m}_i^1 = p_1 F_i^{-1}(\alpha_1) + p_2 F_i^{-1}(0.5) + p_3 F_i^{-1}(\alpha_3), \quad i = 1, \ldots, n,$$

Figure 5 Errors in the Variance for Distribution-Specific Discretization Methods



$p_1 + p_2 + p_3 = 1,$

$p_3 = p_1,$

$\alpha_3 = 1 - \alpha_1,$

$0 \le p_1,\ p_2,\ p_3 \le 1.$

This procedure yields a symmetric discretization, i.e., where $p_3 = p_1$ and $\alpha_3 = 1 - \alpha_1$. LLH found that for $\alpha_1 = 0.05$, the best three-point mean-approximation formula is (P5, P50, P95, 0.179, 0.642, 0.179). This shortcut has very similar probabilities to EPT (P5, P50,

Table 7 Distribution-Specific Method Errors for Main Pearson Types (Discretized Minus Actual Mean and Variance)

	Mean			Variance		
	AE	ASE	ME	AE	ASE	ME
I-U (Beta)						
BMd3	−0.020	9.23E−03	−0.335	−0.288	1.78E−01	−0.982
BMd5	0.004	1.45E−03	0.156	−0.007	1.82E−02	0.371
BMn3	0.000	1.18E−10	0.000	−0.207	4.90E−02	−0.446
BMn5	0.000	1.19E−10	0.000	−0.103	1.29E−02	−0.262
I-J (Beta)						
BMd3	−0.097	9.84E−03	−0.121	−0.466	2.27E−01	−0.616
BMd5	−0.051	2.74E−03	−0.068	−0.284	8.73E−02	−0.395
BMn3	0.000	8.63E−12	0.000	−0.248	6.27E−02	−0.299
BMn5	0.000	7.62E−12	0.000	−0.156	2.58E−02	−0.212
I-∩ (Beta)						
BMd3	−0.068	5.35E−03	−0.108	−0.404	1.69E−01	−0.538
BMd5	−0.040	1.88E−03	−0.066	−0.260	7.21E−02	−0.384
BMn3	0.000	8.69E−14	0.000	−0.222	5.04E−02	−0.292
BMn5	0.000	4.48E−14	0.000	−0.130	1.84E−02	−0.205
VI (Beta Prime)						
BMd3	−0.087	7.74E−03	−0.111	−0.504	2.55E−01	−0.549
BMd5	−0.055	3.06E−03	−0.068	−0.357	1.29E−01	−0.401
BMn3	0.000	4.01E−14	0.000	−0.271	7.37E−02	−0.309
BMn5	0.000	1.69E−14	0.000	−0.185	3.48E−02	−0.230
IV						
BMd3	−0.042	2.20E−03	0.084	−0.508	2.59E−01	−0.548
BMd5	−0.028	9.54E−04	0.055	−0.366	1.35E−01	−0.409
BMn3	0.000	6.34E−09	0.000	−0.262	6.87E−02	−0.287
BMn5	0.000	3.47E−12	0.000	−0.180	3.28E−02	−0.205

Note. Shaded areas are the most accurate distribution-specific methods for specific Pearson types.

P95, 0.185, 0.630, 0.185), which also was constructed using points in similar areas of the (β_1, β_2) plane. However, LLH more densely covered this area, particularly the Type IV region.

5.1. EPT Extensions

Both Pearson and Tukey (1965) and LLH considered a limited set of symmetric discretizations. LLH considered four sets of percentiles ($\alpha_1 = 0.01, 0.05, 0.10, 0.25$) based on the probability elicitation literature. Pearson and Tukey (1965) investigated a different set of percentiles ($\alpha_1 = 0.005, 0.01, 0.025, 0.05$). In this section, we vary α_1 over a larger, more complete set of percentiles, from P1 to P20 in increments of 1%, first maintaining that the discretization must be symmetric, and then relaxing this requirement. Because these shortcuts extend Pearson and Tukey (1965), we refer to them as "EPT+" discretizations and add an identifier that specifies the Pearson region for which a specific discretization is optimized.

For each set of percentiles, we find the probabilities that minimize the ASE. The error-minimizing shortcut discretizations for the Type I, IV, and VI distributions are given in Table 8 (please see the online supplement for the Type III and Type V EPT+ shortcuts). For some of these shortcuts, as well as those in the tables presented later, the fitted probabilities did not sum to one after rounding to three decimal places, but were minimally adjusted to do so. The adjustment was accomplished by successively adding 0.001 to the rounded probabilities in the descending order of the amount lost to rounding (if the probability was rounded down) until the probabilities summed to one. These shortcut methods are new, both in the procedure we use to find them and by their tailoring to specific types of distributions within the Pearson system.

In Table 8, the EPT1∩+ shortcut uses the same percentiles and probabilities (rounded to three decimal places) as EPT. This is due, in part, to Pearson and

Table 8 Symmetric Type-Specific Discretization Shortcuts (EPT+ Methods)

Distribution type	Percentile points	Respective probabilities
I-U Beta (EPT1U+)	P15, P50, P85	0.296, 0.408, 0.296
I-J Beta (EPT1J+)	P6, P50, P94	0.203, 0.594, 0.203
I-∩ Beta (EPT1∩+)	P5, P50, P95	0.184, 0.632, 0.184
VI Beta Prime (EPT6+)	P4, P50, P96	0.164, 0.672, 0.164
IV (EPT4+)	P6, P50, P94	0.212, 0.576, 0.212

Table 9 Nonsymmetric Type-Specific Discretization Shortcuts (EPT++ Methods)

Distribution type	Percentile points	Respective probabilities
I-U Beta (EPT1U++)	P1, P50, P85	0.216, 0.491, 0.293
I-J Beta (EPT1J++)	P2, P50, P94	0.184, 0.615, 0.201
I-∩ Beta (EPT1∩++)	P5, P50, P95	0.184, 0.632, 0.184
VI Beta Prime (EPT6++)	P4, P50, P96	0.164, 0.672, 0.164
IV (EPT4++)	P7, P50, P94	0.231, 0.551, 0.218

Tukey's (1965) heavy sampling of this region. EPT6+ uses the same percentiles and almost exactly the same probabilities as ZDI, perhaps because these distributions are unbounded above, and ZDI is the Gaussian quadrature for the normal distribution. The EPT1U+ and EPT4+ shortcuts do not resemble those of any of the preexisting shortcuts that we consider, although EPT1U+ uses almost exactly the same probabilities as ESM, and the EPT4+ probabilities are similar to those of the MCS. Plots of the absolute errors in the mean and variance for each of the shortcuts discussed in §5 are included in the online supplement.

We now relax the constraint that the discretizations must be symmetric, but we still require that one point be the P50. We again consider values for the lower (upper) percentiles from P1 (P99) to P20 (P80) in increments of 1%. The ASE-minimizing discretizations, which we refer to as the "EPT++" shortcut methods, are shown in Table 9 (please see the online supplement for the Type III and Type V EPT++ shortcuts). The Type I-∩ and VI shortcuts have nearly the same percentiles and probabilities as the symmetric shortcuts in Table 8, implying that three points (that include P50) would not better approximate these pdfs. The Type I-J and IV shortcuts are similar to their symmetric counterparts in their percentiles and probabilities. Only Type I-U is significantly altered by allowing for asymmetry. Thus, it appears that very little accuracy will be gained without increasing the number of points or allowing the middle point to change from P50.

The skewed distributions we consider in this section all have positive skewness, which results in discretizations with more extreme upper percentiles for some distribution types (i.e., the upper percentile α_3 is farther from the median than is the lower percentile α_1). The upper percentile of the Type IV shortcut, for example, is slightly more extreme than the lower percentile, and corresponds specifically to the "thicker" upper tail. If the distribution is left skewed, then the lower percentile should be more extreme to match that tail. Therefore, the shortcuts will need to be accordingly reflected for left-skewed distributions. For example, the shortcut for a right-skewed Type IV is (P7, P50, P94, 0.231, 0.551, 0.218), but for a left-skewed Type IV, the shortcut would become (P6, P50, P93, 0.218, 0.551, 0.231).

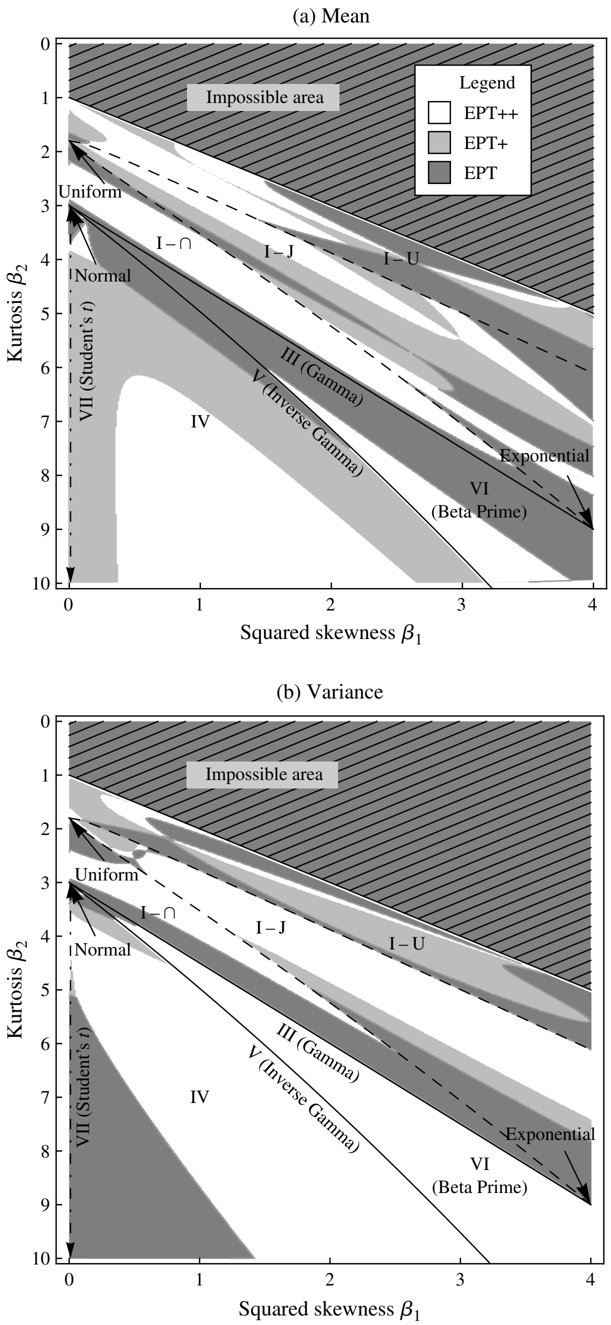
Figure 6 shows which of EPT, EPT+, and EPT++ best matches the mean or variance for specific distributions. Tailoring EPT+ and EPT++ to specific distribution types in most cases improves performance in matching the mean. No one shortcut dominates the others over any entire region, and the shortcut most accurate in the mean of a distribution is often not the most accurate for that distribution's variance.

This improvement in the mean, but in some cases degraded performance in the variance, is a result of our procedure considering only error in matching the means of a set of distributions, without consideration of the variance. One could consider some weighting of the mean and the variance to develop other approximations. Indeed, Keefer's (1994) study considers all the moments via the computation of a certain equivalent. However, developing a shortcut method that preserved certain equivalents would require knowledge of the decision maker's utility function, which is likely to differ widely across decision makers and decision situations.

5.2. Standard Percentile Discretizations

The P10, P50, and P90 percentiles have become so common in practice that displacing them with other percentiles may be difficult. For example, the two most common shortcuts, ESM and MCS, both use these percentiles. Therefore, in this section, we find the best three-point discretizations using these

Figure 6 Regions Where EPT, EPT+, or EPT++ Is the Most Accurate



percentiles. The ASE-minimizing discretizations are given in Table 10. We refer to these as the standard-percentile (SP) shortcut methods, plus identifiers for the regions to which they correspond. Please see the online supplement for the Type III and Type V SP discretizations.

Table 10 Type-Specific P10–P50–P90 Discretization Shortcuts (SP Methods)

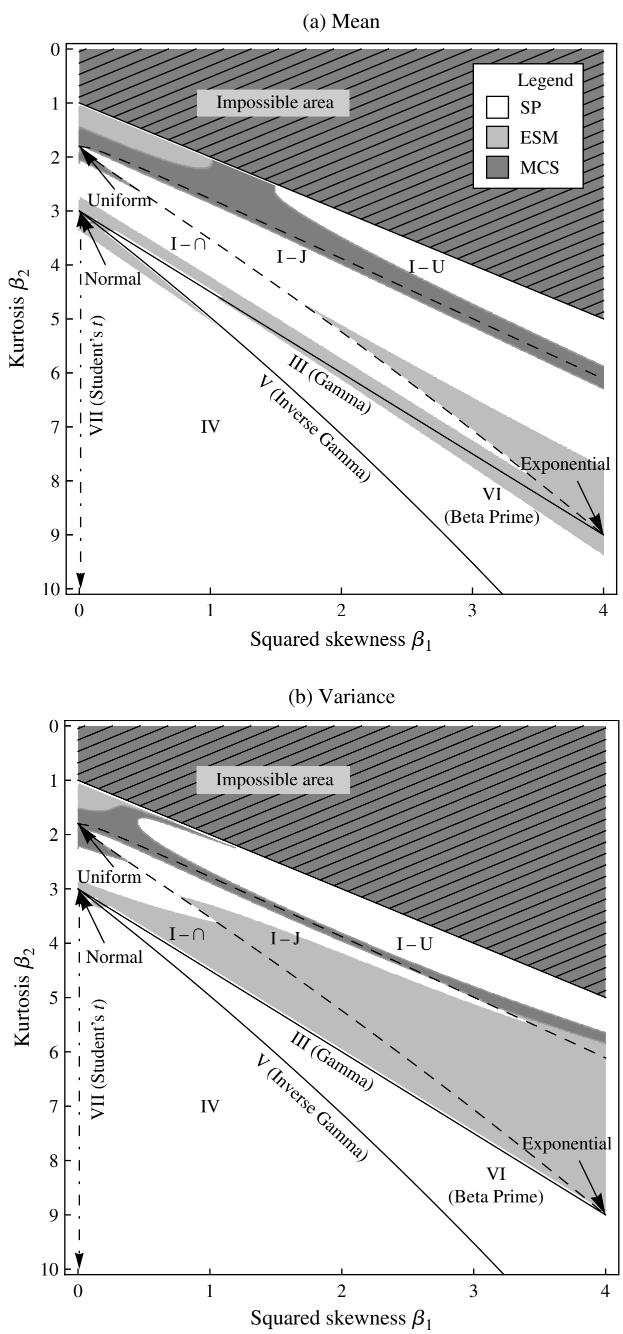
Distribution type	Points	Respective probabilities
I-U Beta (SP1U)	P10, P50, P90	0.228, 0.544, 0.228
I-J Beta (SP1J)	P10, P50, P90	0.273, 0.454, 0.273
I-∩ Beta (SP1∩)	P10, P50, P90	0.296, 0.408, 0.296
VI Beta Prime (SP6)	P10, P50, P90	0.308, 0.384, 0.308
IV (SP4)	P10, P50, P90	0.322, 0.356, 0.322

The probabilities for SP1∩ and SP6 are similar to those of ESM (0.300, 0.400, 0.300), lending previously unknown support to the ESM shortcut; i.e., if one wants to use the P10, P50, and P90 percentiles and is dealing with a unimodal distribution bounded at one or both ends, ESM is close to the discretization that minimizes the error in the mean. However, GQ will still outperform ESM (Bickel et al. 2011). The probabilities for SP1J and SP1U are similar to those of the MCS. The probabilities assigned to the P10 and P90 points are largest for the Type IV distributions because of their higher kurtosis, being unbounded in both directions. The SP4 discretization is almost an equal weighting of the P10, P50, and P90, which is similar to ZDT.

Figure 7 shows which of SP, ESM, and MCS methods, all using the same percentiles, is most accurate in estimating the mean or variance for a given distribution. Our new SP discretizations are the best of the three over most of the region considered, and are the most accurate for both the mean and the variance over nearly the entire Type IV and VI regions. The SP method is a distinct improvement over the ESM and MCS. The MCS is only better than the other two for parts of the Type I-U region. ESM is most accurate in the mean only for small portions of the Type I-J and I-∩ regions, but is most accurate in the variance over most of these same regions.

We now compare the errors for the SP shortcuts to the ESM and MCS. These results are summarized in Table 11, with results for ESM and MCS reproduced from Table 6 and the best error measures highlighted. Errors for these shortcuts for Types III and V are included in the online supplement. The SP shortcuts outperform ESM and MCS, when averaged over the region of interest, in terms of matching the mean, because they use the same percentiles, but fitted probabilities. Indeed, the three error measures

Figure 7 Regions Where ESM, MCS, or SP Is the Most Accurate



for SPs in each region are equal to, or smaller than, those for the ESM and MCS (except for the ME in the Type I-U mean, which is higher than that for the MCS). In terms of the variance, they are more accurate than the MCS by each measure in each region, and generally perform better than the ESM as well, with Type I-∩ and I-J distributions as the only exceptions. SP1∩ is worse than ESM by each measure for Types I-J and I-∩, but better by each measure than the MCS. It is also interesting that SP1U is the best of all shortcut methods considered in this paper at matching the Type I-U variance, as seen by comparing these results to Table 6.

6. Recommendations and Conclusion

In this paper, we have tested the accuracy of existing discretization methods over a much wider set of distributions than has been done previously and developed several new discretization shortcuts, which are more accurate in many cases than previously proposed methods. We conclude with some observations and recommendations for practice.

If the goal of a discretization is to match the moments of the underlying pdf, GQ is ideal because it perfectly matches the first several moments (depending on the number of points used) of any distribution with finite moments. However, GQ can be difficult to implement in practice. Unless one is discretizing a common family (e.g., normal, uniform, triangular), for which GQs have been tabulated (Bickel et al. 2011), use of GQ requires software. In addition, GQ requires that the moments of the underlying distribution be known. BMn matches the mean and is relatively simple to use with a mathematical software package or the graphical method described by McNamee and Celona (1990, p. 30), the latter of which can generally be effectively applied manually (Bickel et al. 2011). However, BMn can significantly underestimate the variance, and usually preserves it less accurately than some shortcut methods (e.g., EPT). BMn's accuracy can be increased somewhat by using more points. One drawback, however, of both the BMn and BMd methods is that the entire distribution is required, whereas shortcut methods need only specific percentiles. That aside, BMd performs quite poorly and should be avoided if one's goal is to closely approximate a distribution's moments.

If the distribution is not known, and especially if assessments are time consuming, we are left with the shortcut methods. Of the preexisting shortcut methods, EPT is a good general choice, but the EPT+ and EPT++ shortcuts presented here generally

Table 11 Errors for the SP Shortcut Methods Compared to ESM and MCS (Discretized Minus Actual Mean and Variance)

	Mean			Variance		
	AE	ASE	ME	AE	ASE	ME
I-U (Beta)						
SP1U	−0.009	9.65E−03	−0.396	0.080	3.70E−02	−0.509
ESM	0.153	3.64E−02	0.449	0.304	1.29E−01	0.800
MCS	0.040	1.11E−02	0.302	0.155	5.61E−02	0.607
I-J (Beta)						
SP1J	−0.003	2.32E−04	0.037	−0.134	2.62E−02	−0.276
ESM	0.034	1.68E−03	0.096	−0.067	1.39E−02	−0.212
MCS	−0.035	1.36E−03	−0.051	−0.197	4.58E−02	−0.330
I-∩ (Beta)						
SP1∩	0.001	2.16E−05	0.016	−0.067	1.04E−02	−0.213
ESM	0.004	4.06E−05	0.019	−0.053	8.73E−03	−0.200
MCS	−0.030	1.05E−03	−0.050	−0.204	4.53E−02	−0.321
VI (Beta Prime)						
SP6	0.000	9.79E−06	0.007	−0.145	2.29E−02	−0.202
ESM	−0.006	4.46E−05	−0.012	−0.166	2.93E−02	−0.221
MCS	−0.042	1.84E−03	−0.051	−0.298	9.01E−02	−0.341
IV						
SP4	0.000	3.53E−06	0.005	−0.113	1.43E−02	−0.167
ESM	−0.007	6.01E−05	0.012	−0.173	3.12E−02	−0.224
MCS	−0.022	6.09E−04	0.044	−0.309	9.66E−02	−0.353

Note. Shaded areas are most accurate shortcut for specific Pearson types.

improve this performance for their respective distribution types. If one is dealing with a distribution that is thought to have low skew, then EPT+ might be a better choice than EPT++.

In practice, the analyst may not be dealing with a distribution that belongs to the Pearson system. However, Pearson and Tukey (1965) showed that distributions that share the same skewness and kurtosis are often very close in shape, even if they are not contained within the Pearson system. With this knowledge, the appropriate EPT+, EPT++, or SP shortcut can be determined by considering the distribution's support (bounded at both ends, bounded at one end, or unbounded). This basic characteristic is determined by the nature of the uncertain quantity and should be apparent. If the distribution is bounded on both ends, then the analyst may be able to use knowledge of its shape (e.g., is it ∩-shaped?) to further specify the appropriate Type I approximation. If there is sufficient knowledge of the distribution to more narrowly specify its location in Figure 1 (if it indeed falls within this region), then Figure 6 or 7 can be used to recommend a particular shortcut. As a practical matter, the SP shortcuts use well-accepted percentiles, whose assessment may be more reliable. For example, Selvidge (1980) and Davidson and Cooper (1980) reported that the P10 and P90 percentiles are assessed more accurately than the P1 and P99 or the P5 and P95. If so, then our new SP approximations will tend to offer improved performance over the commonly used ESM and MCS.

The Pearson system is composed of smooth distributions, most of which are unimodal, with Type I-U as the exception. This type, and perhaps other oddly shaped or multimodal distributions, should be discretized with care. Neither the preexisting shortcut methods we analyze nor the new shortcuts we present perform well over even a quarter of the Type I-U region, which is strong evidence that general shortcut methods will not accurately represent them. A method that takes the actual distribution into account, such as BMn, is better for these kinds of distributions.

As a general approach, shortcut methods are useful as a first approximation, which, when aided by sensitivity analysis, will help identify important uncertainties. These uncertainties can then be given more attention when ascertaining the full distribution and using discretization methods such as Gaussian quadrature or BMn. Decision analysis is an iterative process, and, as the analysis evolves, the discretizations that are used can and should be adapted to the importance of specific uncertainties. How a distribution is ultimately treated in a decision problem is a function not only of the distribution itself, but also of its relation to other aspects of the problem. It is often the case that the analysis arbitrarily does not consider refinement of uncertainty assessments or discretizations (Bickel et al. 2011), and although refinement is not always necessary, its appropriateness should be determined from characteristics of the decision.

Keefer and Bodily (1983) and Keefer (1994) concluded that EPT is a good general discretization method, and our results extended this conclusion to a much wider range of distribution shapes and for different support ranges. As we show, tailoring three-point discretizations to specific distribution families improves accuracy. Although these methods are the result of minimizing average squared error, they also perform well as judged by other measures of accuracy.

In sum, we hope this paper will provide researchers and practitioners with a better understanding of discretization accuracy and that our newly developed discretizations will enjoy widespread use.

Electronic Companion

An electronic companion to this paper is available as part of the online version at http://dx.doi.org/10.1287/deca.1120.0260.

Acknowledgments

The authors gratefully acknowledge the support of the Center for Petroleum Asset Risk Management and the helpful feedback and comments of the anonymous reviewers and the editor-in-chief.

References

Bickel JE, Lake LW, Lehman J (2011) Discretization, simulation, and Swanson's (inaccurate) mean. *SPE Econom. Management* 3(3):128–140.

Burr I (1973) Parameters for a general system of distributions to match a grid of α_3 and α_4. *Comm. Statist.* 2(1):1–21.

Butterworth N (1987) A fast and flexible family of distributions for simulation studies. *IIE Trans.* 19(4):439–444.

Davidson LB, Cooper DO (1980) Implementing effective risk analysis at Getty Oil Company. *Interfaces* 10(6):62–75.

D'Errico JR, Zaino NA (1988) Statistical tolerancing using a modification of Taguchi's method. *Technometrics* 30(4):397–405.

Elderton WP, Johnson NL (1969) *Systems of Frequency Curves* (Cambridge University Press, London).

Hurst A, Brown GC, Swanson RI (2000) Swanson's 30-40-30 rule. *AAPG Bull.* 84(12):1883–1891.

Johnson D (1998) The robustness of mean and variance approximations in risk analysis. *J. Oper. Res. Soc.* 49(3):253–262.

Johnson D (2002) Triangular approximations for continuous random variables in risk analysis. *J. Oper. Res. Soc.* 53(4):457–467.

Johnson NL (1949) Systems of frequency curves generated by methods of translation. *Biometrika* 36(149):78–82.

Keefer DL (1994) Certainty equivalents for three-point discrete-distribution approximations. *Management Sci.* 40(6):760–773.

Keefer DL, Bodily SE (1983) Three-point approximations for continuous random variables. *Management Sci.* 29(5):595–609.

Keefer DL, Verdini WA (1993) Better estimation of PERT activity time parameters. *Management Sci.* 39(9):1086–1091.

Lau H-S, Lau AH, Ho C-J (1998) Improved moment-estimation formulas using more than three subjective fractiles. *Management Sci.* 44(3):346–351.

McNamee P, Celona J (1990) *Decision Analysis with Supertree*, 2nd ed. (The Scientific Press, South San Francisco).

Megill RE (1977) *An Introduction to Risk Analysis* (Petroleum Publishing Company, Tulsa, OK).

Miller AC III, Rice TR (1983) Discrete approximations of probability distributions. *Management Sci.* 29(3):352–362.

Pearson ES, Tukey JW (1965) Approximate means and standard deviations based on distances between percentage points of frequency curves. *Biometrika* 52(3/4):533–546.

Pearson K (1895) Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material. *Philosophical Trans. Royal Soc. London* 186:343–414.

Pearson K (1901) Mathematical contributions to the theory of evolution. X. Supplement to a memoir on skew variation. *Philosophical Trans. Royal Soc. London* 197:443–459.

Pearson K (1916) Mathematical contributions to the theory of evolution. XIX. Second supplement to a memoir on skew variation. *Philosophical Trans. Royal Soc. London* 216:429–457.

Ramberg JS, Schmeiser BW (1974) An approximate method for generating asymmetric random variables. *Comm. ACM* 17(2):78–82.

Rose PR (2001) *Risk Analysis and Management of Petroleum Exploration Ventures*, Methods in Exploration Series (American Association of Petroleum Geologists, Tulsa, OK).

Schmeizer B, Deutsch S (1977) A versatile four parameter family of probability distributions suitable for simulation. *AIIE Trans.* 9(2):176–181.

Slevidge JE (1980) Assessing the extremes of probability distributions by the fractile method. *Decision Sci.* 11(3):493–502.

Smith JE (1993) Moment methods for decision analysis. *Management Sci.* 39(3):340–358.

Taguchi G (1978) Performance analysis design. *Internat. J. Production Res.* 16(6):521–530.

Zaino NA, D'Errico JR (1989) Optimal discrete approximations for continuous outcomes with applications in decision and risk analysis. *J. Oper. Res. Soc.* 40(4):379–388.

Robert K. Hammond is a Ph.D. student in the operations research program and a member of the Decision Analysis Research Group at the University of Texas at Austin. His research interest is decision analysis, focusing on modeling

and quantification of uncertainty for decision making and valuation. His current research is evaluating and improving methods of discretizing continuous distributions. He holds a B.S. degree in electrical engineering and an M.S. degree in operations research, both from the University of Texas at Austin.

J. Eric Bickel is an assistant professor in the Graduate Program in Operations Research, Department of Mechanical Engineering, at the University of Texas at Austin. In addition, he is a fellow in both the Center for International Energy and Environmental Policy and the Center for Petroleum Asset Risk Management. He holds an M.S. and Ph.D. from the Department of Engineering-Economic Systems at Stanford University. His research interests include the theory and practice of decision analysis and its application in the energy and climate-change arenas. His research has addressed the modeling of probabilistic dependence, value of information, scoring rules, calibration, risk preference, education, decision making in sports, and climate engineering as a response to climate change. Prior to joining the University of Texas at Austin, he was a senior engagement manager for Strategic Decisions Group. He is president-elect for the Decision Analysis Society of INFORMS and an associate editor for *Decision Analysis*.