MANAGEMENT SCIENCE
Vol. 29, No. 5, May 1983
Printed in U.S.A.

THREE-POINT APPROXIMATIONS FOR CONTINUOUS RANDOM VARIABLES*

DONALD L. KEEFER† AND SAMUEL E. BODILY‡

This paper compares a number of approximations used to estimate means and variances of continuous random variables and/or to serve as substitutes for the probability distributions of such variables, with particular emphasis on three-point approximations. Numerical results from estimating means and variances of a set of beta distributions indicate surprisingly large differences in accuracy among approximations in current use, with some of the most popular ones such as the PERT and triangular-density-function approximations faring poorly. A simple new three-point approximation, which is a straightforward extension of earlier work by Pearson and Tukey, outperforms the others significantly in these tests, and also performs well in related multivariate tests involving the Dirichlet family of distributions. It offers an attractive alternative to currently used approximations in a variety of applications.
(PROBABILITY MODELING, APPROXIMATION; DECISION ANALYSIS)

1. Introduction

In many management science applications, it is useful to represent the key uncertainties by cumulative distribution functions (CDF's) defined over continuous random variables, or uncertain quantities. The CDF may be elicited from an "expert" via judgmental assessments [5], [12], [18]–[20] or obtained from historical data [22, Chapter 7]. It is common practice—although not always appropriate—to obtain and use just several points from each of the CDF's, especially if the data for a number of CDF's must be elicited judgmentally from busy managers. With PERT and related techniques, for example, three points are generally assessed for each component activity time in order to estimate its mean and variance, thereby permitting the aggregate completion time for the project to be estimated [9], [21]. In risk analysis (Monte Carlo) simulations, on the other hand, it is popular to elicit three points per distribution and fit a triangular probability density function to these [2], [10], [14]. This trend may be furthered by the rapidly increasing availability of easy-to-use financial planning software packages that have this capability.

Three-point approximations (including some which fit a triangular density function to the three points) are also used to estimate parameters for distributions in a variety of analytical applications contexts [3], [4], [13], [23], [24]. Furthermore, when utilizing decision or probability trees, it is convenient to assess just several points for each distribution at the outset rather than to assess each distribution in detail and subsequently devise discrete approximations. In this case, the discrete distribution serves as a substitute for the entire continuous distribution.

This paper compares the accuracy of a number of approximations used to represent continuous distributions and/or to estimate their parameters. Particular emphasis is placed on approximations requiring just three points from the underlying CDF due to their widespread use [2], [3], [4], [9], [10], [13], [14], [21], [23], [24]. Our basic purpose is two-fold:

(a) To point out the surprisingly large differences in accuracy among approximations in current use.

*Accepted by Robert L. Winkler; received October 26, 1981. This paper has been with the authors 2 weeks for 1 revision.

†Gulf Oil Corporation, Pittsburgh.
‡University of Virginia, on leave at the University of Washington.

0025-1909/83/2905/0595$01.25

(b) To suggest the use of a simple new three-point approximation, which is a straightforward extension of earlier work by Pearson and Tukey.

The comparisons are based primarily on the ability of the various approximations to estimate means and variances. In some applications, this is all that is required as noted above. Moreover, the ability to estimate means and variances seems a reasonable minimum requirement for any approximation being considered for more demanding service—e.g., as a substitute for the CDF as a whole.

We begin by describing the approximations for the means and variances, including several related to more general approximations for the entire CDF. These are compared based on their accuracy when applied to a variety of CDF's from the beta family of distributions. (Portions of this work were reported in [7].) Then the performance of the best several approximations are studied in the multivariate context using members of the Dirichlet family of distributions. Recommendations concerning the use of these approximations are provided, and suggestions for further research are offered.

2. Approximations

This section presents the equations for the approximations studied, along with brief comments on their origin. All are based primarily on empirical observation in estimating parameters for univariate distributions. Those for the mean are generally not related to those for the variance unless stated otherwise. We begin with the approximations for the mean.

The first approximation is the classical PERT formula for the mean [9]:

$$\text{mean} = [x(0.0) + 4x_m + x(1.0)]/6, \tag{1}$$

where x_m and $x(p)$ denote, respectively, the mode and the p fractile of the random variable X. Modifications to (1) that use the 0.05 and 0.95 fractiles rather than the extremes of the distribution have been proposed by Moder and Rodgers [11] and by Perry and Greig [16]:

$$\text{mean} = [x(0.05) + 4x_m + x(0.95)]/6, \tag{2}$$

$$\text{mean} = [x(0.05) + 0.95x_m + x(0.95)]/2.95. \tag{3}$$

More recently, Davidson and Cooper [3] developed an approximation using the 0.10 and 0.90 fractiles and the mode:

$$\text{mean} = [x(0.10) + 2x_m + x(0.90)]/4. \tag{4}$$

Our own work indicates this approximation can be improved by modifying its coefficients as follows:

$$\text{mean} = 0.16x_m + .42[x(0.10) + x(0.90)]. \tag{5}$$

Approximations utilizing the median rather than the mode have been proposed by Pearson and Tukey [15] and by Swanson in Megill [10]:

$$\text{mean} = 0.63x(0.50) + 0.185[x(0.05) + x(0.95)], \tag{6}$$

$$\text{mean} = 0.40x(0.50) + 0.30[x(0.10) + x(0.90)]. \tag{7}$$

Most of the simple approximations for the variance involve two symmetrically located fractiles in the tails of the distribution. The classical PERT formula [9] is

$$\text{variance} = ([x(1.0) - x(0.0)]/6)^2. \tag{8}$$

Alternatives using the 0.05 and 0.95 fractiles rather than the extremes have been proposed by Pearson and Tukey [15] and by Moder and Rodgers [11], respectively:

$$\text{variance} = \left(\left[x(0.95) - x(0.05)\right]/3.25\right)^2, \tag{9}$$

$$\text{variance} = \left(\left[x(0.95) - x(0.05)\right]/3.20\right)^2. \tag{10}$$

Moder and Rodgers also proposed an approximation using the 0.10 and 0.90 fractiles,

$$\text{variance} = \left(\left[x(0.90) - x(0.10)\right]/2.70\right)^2, \tag{11}$$

which was modified slightly by Davidson and Cooper [3]:

$$\text{variance} = \left(\left[x(0.90) - x(0.10)\right]/2.65\right)^2. \tag{12}$$

Pearson and Tukey [15] also suggested an iterative scheme requiring five rather than three fractiles (the 0.025, 0.05, 0.50, 0.95, and 0.975 fractiles). We found that an effective three-point approximation for the variance can be obtained by truncating this procedure so that it requires only the 0.05, 0.50, and 0.95 fractiles and is no longer iterative. The applicability of this approximation is significantly enhanced by eliminating the need for the 0.025 and 0.975 fractiles, especially in light of the difficulty of assessing these points in the tails. Specifically, the variance can be estimated by

$$\text{variance} = \left(\left[x(0.95) - x(0.05)\right]/\left[3.29 - 0.1(\Delta/\sigma_0)^2\right]\right)^2, \tag{13}$$

where $\Delta = x(0.95) + x(0.05) - 2x(0.50)$, and σ_0 is the estimate of the standard deviation from (9).

Although numerous pairs of the foregoing mean and variance approximations utilize the same points from the underlying CDF—e.g., (1) and (8) or (6) and (13)—different weights are used to compute the mean and the variance and therefore none corresponds to a single discrete probability distribution. Consequently, these formulas are not particularly helpful in applications where a discrete distribution is required to substitute for the CDF—e.g., in a probability or decision tree. Ideally, we would like to find an approximation that would perform as well as the best of the foregoing formulas in estimating means and variances and could also serve as a discrete-distribution approximation for the CDF.

A discrete-distribution approximation consists of n outcomes $v_1, v_2, \ldots, v_n$ having probabilities $p_1, p_2, \ldots, p_n$. The mean and variance of the approximation are given by

$$\text{mean} = \sum_{i=1}^{n} p_i v_i, \tag{14a}$$

$$\text{variance} = \sum_{i=1}^{n} \left(p_i v_i^2\right) - (\text{mean})^2. \tag{14b}$$

Again we are interested primarily in the case $n = 3$ where a continuous random variable in a decision tree or probability tree is replaced by a three-fork event node.

Pearson and Tukey [15] observed that, for any strictly increasing function $g(\cdot)$ of x, the fractiles of $g(x)$ are simple transformations of those of X; that is, $g(p) = g(x(p))$. Perry and Greig [16] suggested that a sufficiently robust approximation for the mean such as (6) might also provide good results in estimating expected utilities for most utility functions of interest, since $u(x)$ is basically just another random variable for which the mean is to be estimated. Clearly, if an approximation for the mean does perform well in estimating the expected utility of a distribution over a variety of utility functions, it serves for our purposes as a good approximation for that distribution.

Keefer and Pollock [8] concluded that (6) can serve as a very useful approximation for calculating expected utilities and suggested a Taylor series interpretation to help explain its impressive robustness. Therefore, we propose the "extended Pearson–Tukey" approximation as the first of the discrete-distribution approximations to be studied:

$$\begin{aligned} v_1 &= x(0.05), \quad & p_1 &= 0.185, \\ v_2 &= x(0.50), \quad & p_2 &= 0.630, \\ v_3 &= x(0.95), \quad & p_3 &= 0.185. \end{aligned} \tag{15}$$

Since (7) also involves only fractiles from the CDF, we also propose the "extended Swanson–Megill" approximation:

$$\begin{aligned} v_1 &= x(0.10), \quad & p_1 &= 0.3, \\ v_2 &= x(0.50), \quad & p_2 &= 0.4, \\ v_3 &= x(0.90), \quad & p_3 &= 0.3. \end{aligned} \tag{16}$$

In the "bracket-median" approximation, the probability scale of the CDF is divided into a number of equal intervals, or brackets, and the median of each is assigned the probability of its interval. The error in calculating the mean with this approximation can be substantial if only a few intervals are used [19]. Nevertheless, the five-point equiprobability bracket-median approach is commonly used in practice as a discrete-distribution approximation [22, Chapter 5]. It requires the 0.10, 0.30, 0.50, 0.70, and 0.90 fractiles as indicated below:

$$\begin{aligned} v_1 &= x(0.10), \quad & p_1 &= 0.20, \\ v_2 &= x(0.30), \quad & p_2 &= 0.20, \\ v_3 &= x(0.50), \quad & p_3 &= 0.20, \\ v_4 &= x(0.70), \quad & p_4 &= 0.20, \\ v_5 &= x(0.90), \quad & p_5 &= 0.20. \end{aligned} \tag{17}$$

Brown, Kahr, and Peterson [1] propose two discrete-distribution approximations. The first is a three-point approximation based on the median and the extremes of the underlying distribution as indicated below:

$$\begin{aligned} v_1 &= [3x(0.0) + 5x(0.50)]/8, \quad & p_1 &= 0.25, \\ v_2 &= [x(0.0) + 14x(0.50) + x(1.0)]/16, \quad & p_2 &= 0.50, \\ v_3 &= [5x(0.50) + 3x(1.0)]/8, \quad & p_3 &= 0.25. \end{aligned} \tag{18}$$

The second is a five-point approximation that requires the 0.25 and 0.75 fractiles in addition to the median and the extremes:

$$\begin{aligned} v_2 &= x(0.25), \quad p_2 = 0.5[x(0.50) - x(0.25)]/[x(0.25) - x(0.0)], \\ v_4 &= x(0.75), \quad p_4 = 0.5[x(0.75) - x(0.50)]/[x(1.0) - x(0.75)], \\ v_1 &= x(0.25) - 0.25(1 + 2p_2)[x(0.50) - x(0.0)], \quad p_1 = 0.25(1 - 2p_2), \\ v_5 &= x(0.75) + 0.25(1 + 2p_4)[x(1.0) - x(0.50)], \quad p_5 = 0.25(1 - 2p_4), \\ s &= 0.5[x(0.25) + x(0.50)] + 0.5p_2[x(0.50) - x(0.0)], \\ t &= 0.5[x(0.50) + x(0.75)] + 0.5p_4[x(1.0) - x(0.50)], \\ v_3 &= (sp_1 + tp_5)/p_3, \quad p_3 = p_1 + p_5. \end{aligned} \tag{19}$$

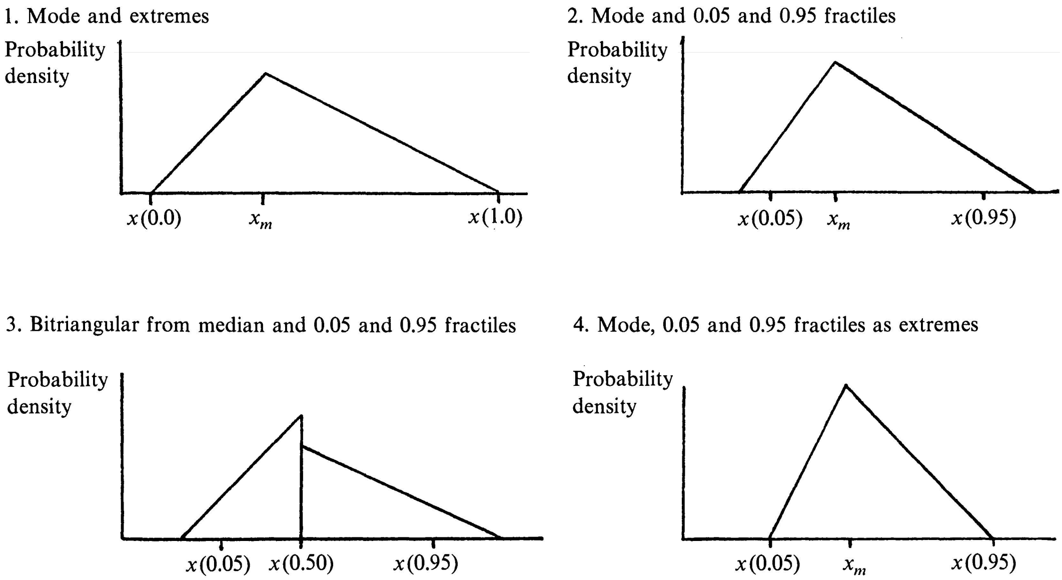


FIGURE 1. Four Three-Point Triangular Models of the Probability Density Function.

For a final philosophy of approximation, we consider triangular probability density functions. These approximations are commonly used in risk analysis (Monte Carlo) simulations [2], [10], [14] and may also be used for computing parameters such as means and variances [13], [23], [24] as noted earlier. The triangular density function generally provides an imperfect representation of the actual density function, but is simple and convenient to use. Its use is often based on the premise that very little information is available about the actual distribution anyway—e.g., only subjective estimates of "high," "low," and "most likely" values. In examining the accuracy of the triangular-based formulas for the mean and variance, we are really evaluating the appropriateness (or quality of fit) of the triangularity assumption over an assortment of probability distributions.

We consider the four triangular models illustrated in Figure 1. The first simply defines the triangular density by lines drawn from the mode to each of the extreme fractiles. The second fits a triangular density to the 0.05 and 0.95 fractiles and the mode. It is defined by determining its extremes, $x(0.0)$ and $x(1.0)$, from two simultaneous nonlinear equations in $x(0.05)$, $x(0.95)$, and x_m. Geometric relationships plus a few lines of algebra give:

$$[x(0.05) - x(0.0)]^2 = 0.05[x(1.0) - x(0.0)][x_m - x(0.0)],$$

$$[x(1.0) - x(0.95)]^2 = 0.05[x(1.0) - x(0.0)][x(1.0) - x_m].$$

These can be solved to find $x(0.0)$ and $x(1.0)$ by standard numerical iterative methods. The third triangular approximation uses the 0.05 and 0.95 fractiles and the median, rather than the mode. It is really a "bi-triangular" approximation, since it results from placing two right triangles back-to-back at the median, with one triangle fit to the median and the 0.05 fractile and the other to the median and the 0.95 fractile. Unless the 0.05 and 0.95 fractiles of the underlying distribution are equidistant from its median, these two triangles will not have the same height—i.e., the bi-triangular density function has a jump, or discontinuity, at the median. The extreme points for the triangles are given directly in terms of the 0.05 (or 0.95) fractiles by:

$$x(0.0) = x(0.50) - [x(0.50) - x(0.05)]/[1 - \sqrt{0.10}],$$

$$x(1.0) = x(0.50) + [x(0.95) - x(0.50)]/[1 - \sqrt{0.10}].$$

The fourth triangular approximation, due to Warren [23], is identical to the first except that it treats the 0.05 and 0.95 fractiles of the underlying distribution as the extremes of the approximating triangular density function, which is fit through these two points and the mode.

3. Numerical Comparisons

3.1. *Method of Comparison*

The numerical comparison of the various approximations is based on their ability to estimate the mean and variance for a set of beta distributions. The standard beta density function with parameters p and q is written:

$$f_x(x) = x^{p-1}(1-x)^{q-1}/B(p,q), \qquad 0 \leqslant x \leqslant 1, \quad p > 0, \quad q > 0, \tag{20}$$

where $B(p,q)$ is the beta function. (If p and q are integers, then

$$B(p,q) = (p-1)!\,(q-1)!/(p+q-1)!.)$$

This family of distributions was chosen for test purposes because it can assume a wide variety of shapes likely to arise in practice. The set of 78 distributions used here corresponds to all combinations of p and q for which $p \leqslant q$, where each can take on the values of 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 30 and 60. It is a modest augmentation of the set of 66 distributions used by Perry and Greig [16] for similar comparisons.

If X is a random variable having a beta distribution as in (20), then its expected value, or mean, and variance are given by

$$\text{mean} = p/(q+p) = p/n, \tag{21a}$$

$$\text{variance} = pq/[(p+q)^2(p+q+1)] = pq/[n^2(n+1)], \tag{21b}$$

where $n = p + q$. The (unique internal) mode is given by $(p-1)/(n-2)$ for the set of distributions considered.

Assuming that the data for the approximations are assessed with perfect accuracy, we can compare the estimates of the mean and variance from the approximations with the actual, or true, values from (21). This simulates the situation where only those points required by each approximation are elicited from the expert (with perfect accuracy), while the type of distribution remains unknown.

Performance is measured in terms of the error and percentage error, which for an actual value z and an approximate value from the formula, $\hat{z}$, are defined as follows:

$$\text{Error} = \hat{z} - z, \tag{22a}$$

$$\%\ \text{Error} = 100(\hat{z} - z)/z. \tag{22b}$$

Most of the results obtained here are easily generalized to beta-distributed variables having the same values of p and q, but bounds other than 0 and 1. For instance, the error in the mean for a beta-distributed variable Y ranging between a and b can be found by multiplying the error for X by $(b - a)$, while the error in the variance for Y is that for X multiplied by $(b - a)^2$. The percentage error in the variance is independent of the range. However, the percentage error in the mean calculated for X applies to Y only as long as the lower bound on Y is zero.

3.2. *Results*

Table 1 shows the maximum and average absolute errors and percentage errors for the various approximations using the test set of beta distributions. Overall, note the wide variation in performances: e.g., the Pearson–Tukey approximation for the mean outperforms the PERT and the first triangular approximation by more than three orders of magnitude on average absolute error. We now examine specific results in more detail, beginning with those for the mean.

3.2.1. *Approximating the Mean.* Clearly, the Perry–Greig, the Pearson–Tukey, the modified Davidson–Cooper, the Swanson–Megill, and the fourth triangular model are superior to the others as approximations for the mean. Note that these approximations are also among the simplest to use of those considered. Despite this result, many of the less desirable formulas apparently are still commonly used: e.g., the PERT-type formulas [21] or the first triangular approximation [2], [10], [13], [14]. Especially poor performances are provided by the original PERT approximation, the Brown–Kahr–Peterson approximations, and the first triangular model. The Pearson–Tukey approximation is extremely accurate, with a maximum percentage error of less than 1/10 of 1% and an average percentage error of about 0.02%.

3.2.2. *Approximating the Variance.* The truncated Pearson–Tukey approximation appears to be the best for estimating the variance, but it is followed very closely by the extended Pearson–Tukey. The latter may be the approximation to use in practice since it is simpler.

There are six other approximations that perform reasonably well: the two-point Pearson–Tukey, the extended Swanson–Megill, the second and third triangular models, the Davidson–Cooper, and the first of the Moder–Rodgers approximations. Particularly poor approximations are the first triangular model (unfortunately perhaps the most commonly used triangular model), the original PERT, and the two Brown–Kahr–Peterson approximations. Again, the differences in performance between the best and the worst are very large. In general the formulas utilizing the 0.05 and 0.95 fractiles (Pearson–Tukey) do somewhat better than those utilizing the 0.10 and 0.90 fractiles.

3.2.3. *Approximating the CDF.* Of those approximations that could be used to substitute for the entire CDF, the clear winner is the extended Pearson–Tukey approximation. The extended Swanson–Megill approximation also does well overall, although not as well for the variance as for the mean. The five-point bracket-median approach performs reasonably well in estimating the mean but poorly in estimating the variance. The first triangular model and both Brown–Kahr–Peterson approximations perform very poorly and should be avoided.

Of the triangular models, the fourth provides the best estimates for the mean, but consistently underestimates the variance by over 50% (by ignoring the tails of the underlying distribution). The bi-triangular model appears to be the best of the four overall, since it is best in estimating the mean while only slightly inferior in estimating the variance. Furthermore it has an advantage from an assessment standpoint because it uses the median—which lends itself to convenient consistency checks—rather than the mode. However, all the triangular models are clearly outperformed by the extended Pearson–Tukey approximation in estimating this set of means and variances.

3.3. *Supplementary Comparisons and Caveats*

The extended Pearson–Tukey (EP-T) approximation appears to be very promising based on these comparisons. Ample evidence of its usefulness in estimating means and variances for univariate distributions is provided in Table 1. It also seems to be a logical choice as a three-point approximation for the entire CDF. In risk analysis simulations where discrete distributions are acceptable (e.g., if the "lumpiness" intro-

TABLE 1

Summary of Errors and Percentage Errors in the Mean and Variance for Beta Distributions

	Approximating the Mean				Approximating the Variance			
	Maximum		Average Absolute		Maximum		Average Absolute	
Approximation (Equation)	Error	% Error	Error	% Error	Error	% Error	Error	% Error
Original PERT (1)	0.14552	451.	0.05224	41.7				
Moder–Rodgers (2)	– 0.04494	– 23.6	0.01148	6.09				
Perry–Greig (3)	– 0.00486	2.05	0.00065	0.37				
Davidson–Cooper (4)	– 0.03499	– 19.1	0.00915	4.88				
Modified Davidson–Cooper (5)	– 0.00139	0.74	0.00024	0.14				
Pearson–Tukey Two-Point (6)	0.00015	0.07	0.00004	0.02				
Swanson–Megill (7)	0.00103	0.33	0.00012	0.05				
Original PERT (8)					0.02728	5506.	0.01768	549.
Pearson–Tukey (9)					0.00085	– 7.5	0.00023	2.2
Moder–Rodgers (10)					0.00201	6.3	0.00055	4.5
Moder–Rodgers (11)					– 0.00167	– 20.7	0.00071	8.8
Davidson–Cooper (12)					0.00271	– 17.7	0.00042	5.6
Pearson–Tukey Truncated (13)					– 0.00086	– 1.7	0.00006	0.38
Extended Pearson–Tukey (15)	0.00015	0.07	0.00004	0.02	– 0.00080	– 1.6	0.00006	0.46
Extended Swanson–Megill (16)	0.00103	0.33	0.00012	0.05	0.00552	11.1	0.00042	2.7
5-Point Bracket Median (17)	– 0.00398	– 3.35	0.00133	0.75	– 0.00551	– 30.2	0.00215	21.5
3-Point Brown–Kahr–Peterson (18)	0.11326	351.1	0.04549	34.65	– 0.03242	4152.	0.01111	365.9
5-Pt. Brown–Kahr–Peterson (19)	0.05510	170.8	0.02076	16.22	0.01362	2634.	0.01019	280.6
Triangular Number 1	0.30663	950.6	0.12675	95.5	0.05415	10928.	0.03456	1040.
Triangular Number 2	0.01267	11.1	0.00420	2.40	– 0.00260	– 7.4	0.00040	3.7
Triangular Number 3	0.00703	4.89	0.00212	1.16	– 0.00260	– 8.7	0.00044	4.1
Triangular Number 4	– 0.00589	– 1.62	0.00055	0.27	– 0.02784	– 55.7	0.00616	54.2

duced is not undesirable), it may be used in place of triangular-density approximations. Similar remarks apply to the extended Swanson–Megill (ES-M) approximation, although its somewhat lower accuracy must be considered. These two approximations do, however, have limitations.

As an illustration, consider a set of lognormal probability distributions. If Y is a normally distributed random variable with mean μ_Y and standard deviation σ_Y, then the random variable X defined by $Y = \log X$ is lognormally distributed. Both the extent to which the density function of X is skewed and the extent to which it is peaked increase rapidly as σ_Y (the standard deviation of the associated normal distribution) increases. Table 2 shows the error and percentage error in the mean and variance of X as a function of σ_Y for the EP-T and ES-M approximations. Note that, although the former significantly outperforms the latter, both approximations deteriorate markedly as σ_Y increases, particularly in their ability to estimate the variance.

These approximations work well for smooth unimodal probability density functions that are not extremely skewed or peaked—the type of function that usually results from judgmental assessment. However, if the distribution appears to be highly skewed or sharply peaked (a suspicion based perhaps on initial assessments) then it is imprudent to use a three-point approximation even to estimate only the mean and variance.

One way to view the quality of the approximation for the "entire distribution" is via the measure of fit used in the Kolmogorov–Smirnov goodness-of-fit test. This is

$$D = \operatorname*{Max}_{x} |F(x) - S(x)| \tag{23}$$

where $F(x)$ is the actual CDF and $S(x)$ is the approximating CDF.

For the approximations using simple fixed fractiles, the measure D will be a constant unaffected by the shape of the underlying distribution. But for those approximations using the mode and for the Brown–Kahr–Peterson approximations, D will depend on the shape of the underlying distribution. Figure 2 shows the determination of D values for the EP-T, ES-M, and five-point bracket median approximations, which turn out to be 0.315, 0.2, and 0.1, respectively. Thus the five-point bracket median approximation has the best "fit" of the three based on this measure.

The triangular distributions would also have low D's. Although the D value for the bi-triangular model depends on the shape of the underlying distribution, it would require a highly unusual distribution to give a D value higher than 0.1.

According to this measure, then, it appears that one of the better triangular approximations (e.g., the bi-triangular) or even the five-point bracket median approximation might be preferable to the EP-T and ES-M approximations if one were primarily concerned with the closeness of the fit of the CDF. However, if this is the major concern, more detailed assessment of the CDF and fitting of a continuous approximating curve may be in order [19].

Finally, we note that there are applications in which a specific interval of a CDF is important. This may occur in decision analysis, for example, if a decision "switches" depending upon whether or not a critical level of an uncertain quantity (e.g., product demand) is reached. In such cases, it is important that this critical interval of the CDF be appropriately represented in a discrete-distribution approximation. Here, it is generally desirable to assess enough points to construct a smooth curve representing the CDF and subsequently to tailor a discrete-distribution approximation to the curve, with particular emphasis on representing the critical interval. Again, a general-purpose three-point approximation cannot always be expected to be appropriate.

Still, if one is looking for simplicity and is not overly concerned with the details of the distribution, the EP-T approximation can be recommended over the alternative simple approximations in current use. Simplicity is often especially important when

TABLE 2

Errors and Percentage Errors for Lognormal Distributions
(Mean of Associated Normal Distribution $\mu_Y = 0.$)

σ_Y		Extended Pearson–Tukey Mean	Extended Pearson–Tukey Variance	Extended Swanson–Megill Mean	Extended Swanson–Megill Variance
0.1	Error	0.00000	− 0.00001	− 0.00008	− 0.00023
	%Error	0.00	− 0.07	− 0.01	− 2.23
0.2	Error	0.00000	− 0.00027	− 0.00038	− 0.00193
	%Error	0.00	− 0.64	− 0.04	− 4.54
0.3	Error	− 0.00006	− 0.00180	− 0.00113	− 0.00855
	%Error	− 0.01	− 1.74	− 0.11	− 8.30
0.4	Error	− 0.00027	− 0.00723	− 0.00271	− 0.02732
	%Error	− 0.03	− 3.55	− 0.25	− 13.42
0.5	Error	− 0.00080	− 0.02284	− 0.00570	− 0.07189
	%Error	− 0.07	− 6.26	− 0.50	− 19.71
0.6	Error	− 0.00192	− 0.06248	− 0.01093	− 0.16755
	%Error	− 0.16	− 10.06	− 0.91	− 26.98
0.7	Error	− 0.00404	− 0.15526	− 0.01956	− 0.36063
	%Error	− 0.32	− 15.04	− 1.53	− 34.94
0.8	Error	− 0.00782	− 0.36068	− 0.03318	− 0.73624
	%Error	− 0.57	− 21.21	− 2.41	− 43.30
0.9	Error	− 0.01421	− 0.79834	− 0.05395	− 1.45169
	%Error	− 0.95	− 28.46	− 3.60	− 51.75
1.0	Error	− 0.02466	− 1.70697	− 0.08477	− 2.80078
	%Error	− 1.50	− 36.55	− 5.14	− 59.96
1.1	Error	− 0.04127	− 3.56312	− 0.12956	− 5.34013
	%Error	− 2.25	− 45.15	− 7.07	− 67.66
1.2	Error	− 0.06707	− 7.32412	− 0.19359	− 10.14279
	%Error	− 3.26	− 53.88	− 9.42	− 74.61
1.3	Error	− 0.10643	− 14.93459	− 0.28396	− 19.31998
	%Error	− 4.57	− 62.35	− 12.20	− 80.66
1.4	Error	− 0.16557	− 30.40591	− 0.41023	− 37.12212
	%Error	− 6.21	− 70.22	− 15.40	− 85.73
1.5	Error	− 0.25331	− 62.17219	− 0.58528	− 72.32489
	%Error	− 8.22	− 77.20	− 19.00	− 89.81

there are many uncertain quantities, particularly when they interact. Indeed, it is in such multivariate contexts that the approximations considered here are most widely utilized. Use of the EP-T approximation and others in this context is explored in the next section.

4. Multivariate Applications

Most applications of approximations for the mean, the variance, or for the entire CDF involve more than one random variable. Specifically, the multivariate problem is

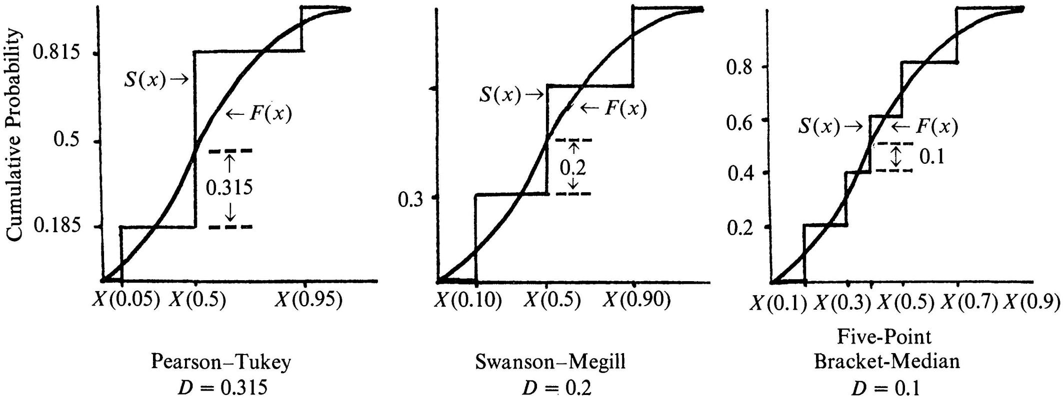


FIGURE 2. Derivation of the Kolmogorov–Smirnov Measure for Three Approximations.

to obtain the probability distribution of the variable, $V = h(X_1, X_2, \ldots, X_n)$, where X_i, $i = 1, 2, \ldots, n$ are random variables. For example, V may be a payoff variable computed from random revenue and cost variables using standard accounting relationships. Unless h is a particularly simple function, the distribution of V may not be easy to determine analytically. Consequently, approximations (or Monte Carlo simulation) play an important role in practical applications. As in the univariate case, it is often sufficient—and generally easier—to calculate parameters such as the mean and variance of V.

4.1. *First Considerations*

The approximation approach we examine replaces a probability or decision tree with continuous event nodes by one with discrete fractiles, for example as shown in Figure 3. In Figure 3, the EP-T approximation is used for each random variable; thus three fractiles of X_1 are assessed (the 0.05, 0.50, and 0.95), then three fractiles of X_2 conditional on X_1 are assessed for each outcome of X_1, and so on ending with the fractiles of X_n, conditional on values $X_1, X_2, \ldots, X_{n-1}$. Clearly 3^n assessments are needed in all (although we give suggestions for short-cuts later). If the X_i's are probabilistically independent, the conditioning is unnecessary and only $3n$ fractiles would have to be assessed to generate the 3^n endpoints. In the general situation, with some independent and some dependent random variables, the order in which the random variables appear in the tree could be chosen so that variables dependent on other random variables would appear to the right of them in the tree.

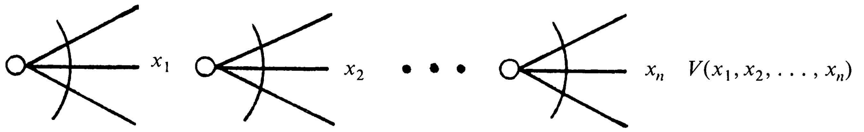
Any of the three-point approximations could be used in this fashion in multivariate applications (or the tree could be expanded to use a five-point approximation). One would expect, however, that the best univariate approximations would work best in the multivariate situation as well. This is obvious in some situations—e.g., if the X_i's are independent and can thus be treated individually and h is a simple function such as a sum or product. More generally, we note that the expectation (or variance) of V can be decomposed into a sequence of one-dimensional expectations (or variance computations). For example the expectation of $V = h(X, Y)$ can be written as

$$E[V] = \int_X E_Y[h(Y \mid x)] f_x(x)\, dx \quad \text{where} \tag{24a}$$

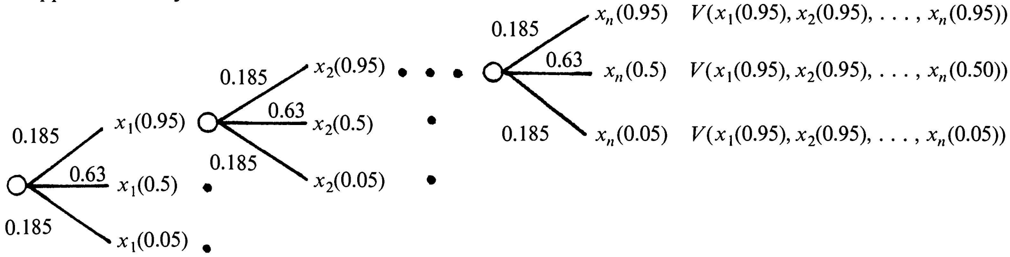
$$E_Y[h(Y \mid x)] = \int_Y h(x, y) f_{Y \mid X}(y \mid x)\, dy. \tag{24b}$$

The stagewise approach described above involves three EP-T approximations of (24b) at $X = x(0.05)$, $x(0.50)$, and $x(0.95)$, respectively, together with an EP-T approximation of (24a) using these results. The point is that if the individual EP-T approxima-

A general probability tree with n random variables,



is approximated by



V is then approximated by 3^n outcomes having probabilities $(0.63)^m$ $(0.185)^{n-m}$ where m is the number of variables set at their median.

FIGURE 3. An Approach to a Multivariate Approximation Using the Extended Pearson–Tukey Univariate Approximation.

tions are good enough, the overall approximation will also be good. These arguments also hold if V is a utility function expressing preferences for payoff levels.

The prime candidates for use in multivariate applications are therefore the EP-T and the ES-M approximations. From the comments in §3.3, of course, some caution is called for in applying the EP-T or ES-M to approximate the CDF of V or even its mean and variance. Moreover, even if V is a function of independent random variables and only the mean of V is being estimated, there will be cumulative effects due to combining the errors in the individual approximations, even though $E[V] = h(E(X_1), E(X_2), \ldots, E(X_n))$. For highly nonlinear h, these errors will be more pronounced. Without probabilistic independence, the interactions between variables may exaggerate the errors even further. Our purpose in this section is to obtain preliminary indications of the quality of approximations involving more than one random variable.

4.2. *Multivariate Numerical Comparisons*

As a first step towards investigating the general multivariate case, we present some results for the bivariate case: i.e., $V = h(X, Y)$. In particular, we consider two of the most common forms for $h(\cdot, \cdot)$, the sum and the product forms:

$$h(X, Y) = X + Y, \tag{25a}$$

$$h(X, Y) = XY. \tag{25b}$$

For each form and for each of the bivariate test distributions, we examine the performance of the various approximations in estimating the mean and variance of V and the covariance of X and Y. These particular terms are important in themselves, but the covariance term takes on added importance in light of its role in approximating the means and variance of more complex nonlinear functions with many variables.

Howard [6] has suggested the following approximations for the mean and variance of V:

$$\text{mean}(V) \simeq h(E(X_1), E(X_2), \ldots, E(X_n)) + \frac{1}{2} \sum_{i,j} \frac{\partial^2 h}{\partial X_i \partial X_j} \text{cov}(X_i, X_j), \tag{26a}$$

$$\text{var}(V) \simeq \sum_{i,j} \frac{\partial h}{\partial X_i} \frac{\partial h}{\partial X_j} \text{cov}(X_i, X_j), \tag{26b}$$

where all derivatives are evaluated at the means of the variables, and $\text{cov}(X_i, X_i) = \text{Var}(X_i)$. Assuming that the means and variances of each X_i can be approximated accurately enough, the usefulness of these approximations will depend on the accuracy in approximating the covariance and the appropriateness of Howard's quadratic approximations (26) in the specific situation.

One reason for using Howard's approximations rather than computing the mean and variance from the complete probability tree is that it can reduce the number of assessments and computation necessary when n is large and probabilistic dependence is present. An estimate of covariance is needed for each pair of variables. Using three-point approximations, nine fractiles are needed for each of the $n(n-1)/2$ pairs of variables. These same fractiles could also provide the approximations for the mean and variance of each X_i except one. The total number of fractiles assessed, $4.5n(n-1)+3$, will be less than the 3^n assessments for the complete tree when $n > 3$. For example, with 6 random variables, the complete tree requires 729 assessments while using a three-point covariance approximation in Howard's expressions requires 138 separate fractiles.

Howard [6] suggests a procedure for estimating covariances to be used in his approximations that is based on drawing curves representing the relationship of the mean of X_j conditional on X_i values. We submit that a decision maker will find our method of providing conditional fractiles easier (he does not have to supply the whole curve), that he will be able to do it more consistently (consistency checks can be incorporated into the fractile assessment), and that he will have greater confidence in his answers. Indeed, if it were just as easy for decision makers to provide means as to provide the median or other fractiles, there would be no need for three-point approximations for the mean at all.

The approximations studied for estimating $E(V)$, $\text{var}(V)$ and $\text{cov}(X, Y)$ were the best of the three-point approximations, i.e. the EP-T and ES-M approximations. In addition, the five-point bracket median approach was included for comparison.

The performances of the approximations were compared using the multivariate analog of the beta distribution, the Dirichlet distribution. The Dirichlet density function has three parameters: $p_x > 0$, $p_y > 0$, and $q > 0$. It is given by

$$f(x, y) = \left[\Gamma(p_x + p_y + q)/(\Gamma(p_x)\Gamma(p_y)\Gamma(q))\right]x^{p_x - 1}y^{p_y - 1}(1 - x - y)^{q-1} \tag{27}$$

where $\Gamma(\cdot)$ is the gamma function. The parameters p_x, p_y, and q were allowed to range separately over the values 2, 3, 6, 10, 20, and 60, giving 216 different bivariate distributions.

Exact expressions for $E(X+Y)$, $\text{var}(X+Y)$, $E(XY)$, $\text{var}(XY)$ and $\text{cov}(X, Y)$ are available in terms of p_x, p_y, and q for use in calculating errors (although for space reasons we will not give the expression here) [17]. Errors have been averaged over the 216 different Dirichlet distributions.

4.3. *Multivariate Results*

Table 3 summarizes the results. Of the approximations, only the EP-T does well for all of the factors. The errors with ES-M are tolerable for many uses but are higher for the variances and the covariance than with the EP-T. The five-by-five bracket-median, as anticipated from the results in Table 1, misses on the variances and the covariance; there seems to be little reason to use it rather than one of the other two.

The error in the $\text{cov}(X, Y)$ using the EP-T is so low (and the errors in means and variances so low from §2) that it seems reasonable to use it for large n to get the terms in Howard's approximations (26). Again, an important consideration is the accuracy of the quadratic approximation of (26) itself. These results are also, of course, encouraging for those situations where the more direct probability tree approach of Figure 3 is

TABLE 3

Errors and Percentage Errors in the Means and Variances of the Sum and Product of Two Uncertain Quantities and in Their Covariance

		Average Absolute Error	Average Error	Average Absolute Percentage Error
Five by Five Bracket Median	mean$(X + Y)$	0.0012955	− 0.0004114	0.35139
	var$(X + Y)$	0.0011715	− 0.0011715	22.40927
	mean(XY)	0.0006014	0.0002503	1.18503
	var(XY)	0.0001295	− 0.0001295	21.42346
	cov(X, Y)	0.0006207	0.0006207	23.74023
Three by Three EP-T	mean$(X + Y)$	0.0000322	0.0000102	0.00764
	var$(X + Y)$	0.0000259	0.0000158	0.40679
	mean(XY)	0.0000177	− 0.0000045	0.03297
	var(XY)	0.0000054	0.0000003	0.57851
	cov(X, Y)	0.0000141	− 0.0000124	0.43388
Three by Three ES-M	mean$(X + Y)$	0.0001109	− 0.0000269	0.02139
	var$(X + Y)$	0.0001907	− 0.0001202	2.78590
	mean(XY)	0.0000832	0.0000094	0.12581
	var(XY)	0.0000363	− 0.0000280	3.95458
	cov(X, Y)	0.0000686	0.0000230	2.79328

employed. This approach would seem the logical choice if n is relatively small or if the variables are independent.

4.4. *Further Research*

This section has only provided a start on the problem of approximating the moments in multivariate applications. But the results for the EP-T approximation are encouraging. It would not be difficult to evaluate errors in approximations for polynomial functions in more than one variable. Similarly, performance of various approximations in estimating higher order moments could be investigated. The results here give a very rough idea, however, of how these explorations will turn out.

It may be useful to develop more specific methods for approximating certain multivariate probability distributions. It may also be possible to develop a better approximation to $f(x, y)$, the joint probability distribution in X and Y, by appropriate selection of fewer than nine points in $X \times Y$ space. Any work on this idea, however, should reflect concerns for how the points may be assessed. People are able to assess the fractiles of $Y | X$; would they be able to assess selected fractiles of the joint distribution of X and Y? Since covariance (or correlation) plays such a key role in estimating the central moments of a function of uncertain quantities, another useful development would be approximations based on limited assessment data developed specifically for estimating the covariance (or correlation).

5. Conclusions

The results reported here indicate the EP-T approximation is a widely-applicable general-purpose three-point approximation for continuous probability distributions. It is more accurate—often by a wide margin—than its competitors in estimating means and variances of distributions typical of those elicited via judgmental assessments. Since the approximation itself is a discrete distribution, it can be used to substitute for continuous distributions where appropriate.

From the probability assessment standpoint, it has the advantage of utilizing the median, which naturally lends itself to straightforward consistency checks, rather than the mode. On the other hand, it does require the 0.05 and 0.95 fractiles (as do many of its competitors), which are somewhat more difficult to assess accurately than points closer to the center of the distribution. The second-best approximation studied here, the ES-M approximation, may be an attractive alternative if this is a major concern, since it utilizes the 0.10 and 0.90 fractiles in addition to the median.

As emphasized throughout, such approximations definitely do have limitations. If three-point approximations are to be used, however, it makes sense to use the best ones available. At the very least, it would be prudent to avoid using those that perform very poorly in simple tasks such as estimating means and variances.

References

1. BROWN, R. V., KAHR, A. S. AND PETERSON, C., *Decision Analysis for the Manager*, Holt, Rinehart, and Winston, New York, 1974.
2. CAIN, WILLIAM O., JR., "Risk Analysis Enhances Exploration," *The Oil and Gas. J.* (November 17, 1980), pp. 119–132.
3. DAVIDSON, L. B. AND COOPER, D. O., "A Simple Way of Developing a Probability Distribution of Present Value," *J. Petroleum Tech.* (September 1976), pp. 1069–1078.
4. ——— AND ———, "Implementing Effective Risk Analysis at Getty Oil Company," *Interfaces*, Vol. 10, No. 6 (1980), pp. 62–75.
5. HOLLOWAY, CHARLES A., *Decision Making Under Uncertainty: Models and Choices*, Prentice-Hall, Englewood Cliffs, N.J., 1979.
6. HOWARD, RONALD A., "Proximal Decision Analysis," *Management Sci.*, Vol. 17, No. 9 (1971), pp. 507–541.
7. KEEFER, D. L., "Three-Point Approximations for Judgmental Probability Distributions," presented at TIMS/ORSA Meeting, Washington, D.C., May 1980.
8. ——— AND POLLOCK, S. M., "Approximations and Sensitivity in Multiobjective Resource Allocation," *Oper. Res.*, Vol. 28, No. 1 (1980), pp. 114–128.
9. MALCOLM, D. G., ROSEBOOM, J. H., CLARK, C. E. AND FRAZAR, W., "Application of a Technique for Research and Development Program Evaluation," *Oper. Res.*, Vol. 7, No. 5 (1959), pp. 646–669.
10. MEGILL, ROBERT E., *An Introduction to Risk Analysis*, Petroleum Publishing Company, Tulsa, 1977.
11. MODER, J. J. AND RODGERS, E. G., "Judgment Estimates of the Moments of PERT Type Distributions," *Management Sci.*, Vol. 15, No. 2 (1968), pp. B76–B83.
12. MOORE, P. G. AND THOMAS, H., *The Anatomy of Decisions*, Penguin Books, New York, 1976.
13. NEELY, WALTER P., SELLERS, JACKIE AND NORTH, RONALD M., "Goal Programming Priority Sensitivity Analysis: An Application in Natural Resource Decision Making Processes," *Interfaces*, Vol. 10, No. 5 (1980), pp. 83–89.
14. NEWENDORP, PAUL D., *Decision Analysis for Petroleum Exploration*, Petroleum Publishing Company, Tulsa, 1975.
15. PEARSON, E. S. AND TUKEY, J. W., "Approximate Means and Standard Deviations Based on Distances Between Percentage Points of Frequency Curves," *Biometrika*, Vol. 52, No. 3–4 (1965), pp. 533–546.
16. PERRY C. AND GREIG, I. D., "Estimating the Mean and Variance of Subjective Distributions in PERT and Decision Analysis," *Management Sci.*, Vol. 21, No. 12 (1975), pp. 1477–1480.
17. PRESS, S. JAMES, *Applied Multivariate Analysis*, Holt, Rinehart and Winston, New York, 1972, pp. 133–135.
18. RAIFFA, HOWARD, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*, Addison-Wesley, Reading, Mass., 1968.
19. SCHLAIFER, ROBERT, *Analysis of Decisions Under Uncertainty*, McGraw-Hill, New York, 1969.
20. SPETZLER, CARL S. AND STAEL VON HOLSTEIN, CARL-AXEL S., "Probability Encoding in Decision Analysis," *Management Sci.*, Vol. 22, No. 3 (1975), pp. 340–358.
21. TAYLOR, B. W. AND MOORE, R. J., "R & D Project Planning with *Q*-GERT Network Modeling and Simulation," *Management Sci.*, Vol. 26, No. 1 (1980), pp. 44–59.
22. VATTER, PAUL A., BRADLEY, STEPHEN P., FREY, SHERWOOD C., JR. AND JACKSON, BARBARA B., *Quantitative Methods in Management*, Richard Irwin, Homewood, Ill., 1978.
23. WARREN, JOSEPH E., "Basin Evaluation," Society of Petroleum Engineers Paper 7717, presented at SPE Economics and Evaluation Symposium, Dallas, February 1979.
24. ———, "The Development Decision: Frontier Areas," Society of Petroleum Engineers Paper 9558, presented at SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, February 1981.