STAFF PAPER

DEPARTMENT OF
AGRICULTURAL ECONOMICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907

Discrete Approximations of Joint Probability Distributions

Eric DeVuyst[1]
Paul V. Preckel

Department of Agricultural Economics
Purdue University

Staff Paper #91-1
January 1991

Abstract

Practical computational limits for stochastic decision analysis models often require that probability distributions have a modest number of points with positive mass. This paper develops an approach to constructing such discrete joint probability distributions which introduces less bias than more commonly used methods. The method, based on solving systems of nonlinear equations, is demonstrated for both continuous and discrete distributions.

Keywords: Decision Analysis, Approximate Probability Distributions, Approximate Joint Probability Distributions.

[1] The authors are graduate assistant and associate professor in the Department of Agricultural Economics at Purdue University, respectively.

1 Introduction

Practical limits to computation for stochastic decision analysis problems often require that probability distributions be represented by a modest number of points with positive mass. Where possible, these distributions should be subjective and elicited directly from the decision maker. When the number of random variables relevant to the decision problem is small and the decision maker has developed intuition regarding the random variables, the direct elicitation of distributional information is reasonably practical. However, when the number of random variables is large or the decision maker has little or no basis for imagining distributions for some or all of the random variables, other methods based on historical observations may be more useful than elicitation.

In situations where historical observations are employed to develop distributional information, two commonly employed approaches are to either directly use the historical data as an "empirical distribution" by assigning probabilities to the individual observations or to use the data to estimate the parameters of a known (or assumed) continuous distribution. Due to practical limits on computational capacity, neither of these approaches may be appropriate since the empirical distribution may contain too many points to be directly useful, and an estimated continuous distribution may not be directly useful. (An example of the latter case occurs in problems of expected utility maximization where, given the continuous distribution of the random variable or variables, the expected utility function has no closed form expression.) In these cases, the method described in this paper may be of use. This method, called Gaussian Quadrature for Joint Distributions, is an approach to approximating joint probability distributions by a set of discrete points and associated probabilities. The method is superior to other commonly used methods which can be shown to systematically understate the variability in the distributions to be approximated.

2 Notation

The notation to be used throughout this paper is defined here. Let:

$x=[x_i]_{i=1,2,\ldots,m}=$ a vector of m random variables whose components are denoted x_i.

$\{x\}=$ the joint probability mass function for the variables x_i, $i=1,2,\ldots,m$.

$<f(x)>=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}f(x)\{x\}dx_1dx_2\ldots dx_3=$ the expected value of the function $f(x)$.

${}^jx=[{}^jx_i]_{i=1,2,\ldots,m}=$ the vector in m-dimensional space which represents the j-th point in the discrete approximation to the distribution $\{x\}$.

${}^jp=$ the probability mass associated with ${}^jx$ in the discrete approximation.

$y=$ a vector of decision variables.

3 Stochastic Decision Problems, Taylor Series, and Gaussian Quadrature

In general, a stochastic decision problem may be stated in the form

$$maximize_{y\in\Omega}<f(x|y)> \tag{1}$$

where $f(x|y)$ is a measure of the benefit associated with the choices y when x is realized and whose expected value is to be maximized. (In the following discussion, the argument y in f will be suppressed.) To make the solution of such problems practical when $<f(x)>$ has no closed form solution, the expected value is typically approximated by

$$<f(x)> \approx \sum_{h=1}^{N} {}^{h}p\, f({}^{h}x). \tag{2}$$

Provided that $f(x)$ has a Taylor series expansion (or that $f(x)$ can be approximated by a polynomial over the relevant range for x) then

$$\begin{aligned} f(x) &= f(\bar{x}) + \sum_{i=1}^{m} (x_i - \bar{x}_i) \frac{\partial f(\bar{x})}{\partial x_i} \\ &+ \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{1}{2!} (x_i - \bar{x}_i)(x_j - \bar{x}_j) \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j} \\ &+ \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{3!} (x_i - \bar{x}_i)(x_j - \bar{x}_j)(x_k - \bar{x}_k) \frac{\partial^3 f(\bar{x})}{\partial x_i \partial x_j \partial x_k} + \dots \end{aligned} \tag{3}$$

Thus, the problem objective becomes the integral of an infinite polynomial in the variables x_i.

Combining (1) and (3),

$$\begin{aligned} <f(x)> &= f(\bar{x}) + \sum_{h=1}^{N} {}^{h}p_i \sum_{i=1}^{m} ({}^{h}x_i - \bar{x}_i) \frac{\partial f(\bar{x})}{\partial x_i} \\ &+ \sum_{h=1}^{N} {}^{h}p_i \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{1}{2!} ({}^{h}x_i - \bar{x}_i)({}^{h}x_j - \bar{x}_j) \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j} \\ &+ \sum_{h=1}^{N} {}^{h}p_i \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{3!} ({}^{h}x_i - \bar{x}_i)({}^{h}x_j - \bar{x}_j)({}^{h}x_k - \bar{x}_k) \frac{\partial^3 f(\bar{x})}{\partial x_i \partial x_j \partial x_k} \\ &+ \dots \end{aligned} \tag{4}$$

If the above approximation of $<f(x)>$ is truncated after K terms, then the approximation involves expectations of products of the random variables only up to the K-th order. If all of these moments are correct for the approximate distribution, the truncation error may be expressed as

$$<\epsilon_{K+1}> = \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \cdots \sum_{i_{K+1}=1}^{m} \langle \frac{\prod_{k=1}^{K+1} (x_{i_k} - \bar{x}_{i_k})}{(K+1)!} \frac{\partial^{K+1} f(\theta \bar{x} + (1-\theta)x)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_K}} \rangle \quad (5)$$

for some $0 \le \theta \le 1$. This error will be small if all of the moments of the K plus first order are small or if the expected value of the K plus first order derivatives of the objective are small. In practice these conditions are likely to be difficult or even impossible to check. However, it is intuitively appealing that extremely high order moments of the distribution are unlikely to have a large effect on decision maker choices. Thus, so long as K is chosen to be moderately large, the truncation error seems likely to be small.

In their paper on discrete approximations of distributions of single random variables, Miller and Rice suggest that the discrete distribution should be chosen so as to match as many as possible of the lower order moments of the true distribution. They go on to show that the usual approach of defining a discrete approximation to a distribution by dividing the domain into regions and assigning the probability of falling in the region to the conditional mean in the region biases the even numbered moments downward. Because the direction of the bias can be predicted (at least for even numbered moments), they suggest that the simple approach to defining approximate discrete distributions should be viewed as unacceptable. As an alternative, Miller and Rice suggest using a gaussian quadrature approach to distribution approximation.

There is little evidence in the literature that this gaussian quadrature approach has been widely adopted by decision analysis practitioners. The probable reasons are two-fold. When possible, it is probably more relevant to elicit the decision maker's (DM's) subjective probability distribution directly. Second, the gaussian quadrature approximation procedure, as it appears in the literature, is focused on univariate distributions. However, it is in the multivariate case that the DM would likely be overwhelmed by the task of specifying probabilities associated with the joint outcomes of several random variables. In this case, an analysis based on historical data is typically needed to develop distributional information.

One approach to what is in essence a problem of numerical integration is to use the Monte Carlo approach to computing the expected benefit function. Thus, a uniform random sample over the support region for the continuous random variables would be generated and the product of the density and the benefit function evaluated at each realization would be computed and summed to give an approximation to the integral. While this has been viewed as a good approach to numerical integration, particularly in the multivariate case, it fails to satisfy some important goals. First, the desired distribution approximation should involve a moderate number of points. To get reasonable accuracy from the Monte Carlo approach, a fairly large number of points must be sampled. Second, the Monte Carlo approach does not take full advantage of our knowledge of the joint distribution of the random variables. Third, there does not seem to be an analog to the Monte Carlo approach for discrete distributions. It is noteworthy that with Monte Carlo integration, not even the mean of the distribution is preserved.

As an alternative, we are proposing an approach to constructing approximate distributions which is not sampling based as is a Monte Carlo method. Before considering the details of the theory and practice of the new method, a simpler approach that has been used in previous studies (*e.g.*, Moss) will be described and illustrated.

The analogous approach to the simple method (described by Miller and Rice) of defining a discrete approximation is no longer entirely simple in the multivariate case. For example, one approach to defining a discrete approximation to a distribution which uses a small number of points and preserves the mean of the true distribution would be to divide the support of the random variables into regions defined by planes parallel to all but one of the axes for the random variables which pass through the mean vector. Points could then be defined by the conditional means for the regions, and the mass associated with the points would be equal to the mass associated with the region. Thus, for the standard bivariate normal distribution for instance, the discrete approximation would be the conditional means for each of the four quadrants, and the mass on each of these points would be one quarter.

Based on this simple approach, four-point distributions were constructed for the standard bivariate normal and bivariate multinomial distributions. These approximate distributions appear in Table 1. The percentage errors in the first four central moments for these approximate distributions are displayed in Table 2. Many of the moments (higher than the first) are substantially in error. In particular, all moments involving only even numbered powers of the random variables are substantially understated. Thus, the same sort of predictable bias observed by Miller and Rice in the univariate case also occurs in the multivariate case.

Table 1: Approximate Distributions Based on the Conditional Expectation Approach

Point No.	Bivariate Normal $f(x,y)=\frac{1}{\sqrt{2\pi}}e^{-x^2-y^2}$			Bivariate Multinomial $f(x,y)=\frac{20!\,2^{-20}}{x!\,y!\,(20-x-y)!}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^y$ where $x+y\leq 20$		
	p	x	y	p	x	y
1	0.25	0.1995	0.1995	0.0945	6.6612	6.6612
2	0.25	-0.1995	0.1995	0.2884	3.5789	7.0870
3	0.25	0.1995	-0.1995	0.2884	7.0870	3.5789
4	0.25	-0.1995	-0.1995	0.3288	3.9388	3.9388

The approach Miller and Rice suggest is to base the approximation on gaussian quadrature. With a single random variable, this amounts to choosing a set of probabilities and points satisfying,

$$<x^k>=\int_{-\infty}^{\infty} x^k\{x\}dx=\sum_{i=1}^{N} {}^hp\,{}^hx^k \tag{6}$$

for $k=0,1,...,2N-1$. In the case where x represents a single random variable, this system of equations is straightforward to solve by solving two linear systems and determining all of the roots of a polynomial of order $N-1$. The resulting solution satisfies some convenient properties. For instance, if the moments of the original distribution are finite, then the solution to the system will exist and the values ${}^hx$ will lie within the support of the original distribution.

Table 2: Percentage Error in Central Moments Using Conditional Expectation Approach

Moment	Bivariate Normal	Bivariate Multinomial
$<x>$	0.0	0.0
$<x^2>$	-96.0	-4.5
$<xy>$	0.0	0.7
$<x^3>$	0.0	-10.5
$<x^2y>$	0.0	-3.5
$<x^4>$	-99.9	-17.9
$<x^3y>$	0.0	-9.1
$<x^2y^2>$	-99.8	-9.0

This paper addresses the generalization and solution of this system of equations for the case where a discrete approximate distribution for a set of jointly distributed random variables is desired. The method is equally applicable to both continuous and discrete joint probability distributions. Special approaches to the joint normal case which result in less efficient but more easily computed approximations are also derived.

4 Joint Distribution Approximations Based on Gaussian Quadrature

The generalization of the approach to approximating distributions based on gaussian quadrature to joint distributions is theoretically straightforward. In the multivariate case, the analogous system of equations to solve to match all moments up to the K-th order (including the cross moments) is given by

$$\sum_{j=1}^{J} {}^{j}p \prod_{i=1}^{m} {}^{j}x_i^{h_i} = \langle \prod_{i=1}^{m} x_i^{h_i} \rangle \tag{7}$$

where J denotes the number of points in the discrete approximation, and for all combinations of the h_i's such that $0 \le \sum_i h_i \le K$. Miller and Rice suggest a method for solving this system in the univariate case which does not generalize to the multivariate case. However, advances in software for numerical solution of nonlinear equations make it reasonable to consider solving this system of equations directly. The structure of the nonlinear system has implications for the number of points in the approximation based on the number of random variables and the number of moments to be exactly matched. That is, if there are two random variables, and it is desired to match the moments from the zero-th through the fifth, then the number of conditions on the discrete distribution is equal to one for the zero-th moment, two for the first moments, three for the second moments, four for the third moments, five for the fourth moments, and six for the fifth moments, yielding a total of twenty one conditions. For a bivariate distribution, each point in the discrete distribution contributes three variables to the system of nonlinear equations: the probability of the point, the value for the first random variable and the value for the second random variable. If we choose to satisfy all moments through the fifth order (21), the number of equations and the number of variables can be chosen to be equal by choosing to have seven points in the discrete approximation to the distribution.

If, on the other hand, it is desired to match exactly only the zero-th through the third moments (ten conditions), then there will either be too many conditions or too many variables depending on how many points are chosen. This occurs because the number of moments is not divisible by one plus the number of random variables. When all moments of order K or less are satisfied exactly, the K-th order Taylor series expansion can be used as the objective function for the decision problem, and the error in the approximation of the objective will be the expected value of ϵ_{K+1}.[2]

In the case where the number of points and the number of moments to be matched are not consistent, the resulting system will either be underdetermined if too many points are used in the approximation, or overdetermined if too few points are used. In the case where the system is underdetermined, there will be multiple solutions to the system of equations. Since there will be differences in the higher order moments for these alternative solutions, the choice of approximate distribution may affect the solution to the decision problem. However, if the distribution is approximated with a sufficient number of moments being satisfied, the resulting effect on the solution to the decision problem should be negligible. One way to create a completely determined system is to include additional higher order moments until the number of variables and nonlinear equations is matched. However, even when the nonlinear system has

[2] Alternatively, the expectation of the original benefit function may be used (where the expectation is taken with respect to the approximating distribution). The approximation error will be somewhat different in this case.

a number of variables equal to the number of equations, the resulting system may have multiple solutions or no solutions as will be seen below in our examples. It appears to be difficult to recognize when multiple solutions exist. No solutions will exist when the system of nonlinear equations is dependent and the moments to be matched are not dependent in the same way. For instance, for a bivariate distribution where we wish to match all moments up to the second order, a straightforward counting of the number of equations and variables suggests that two points are adequate to approximate the distribution. However, due to nonlinear dependencies between the equations, it may not be possible to match both of the variances and the covariances using only two points. Similarly, for five point approximation to a bivariate distribution a straightforward counting of variables indicates that it should be possible to exactly satisfy all moments of order four or lower. However, these equations are dependent, and there may be no solution to the system unless the moments themselves are dependent in the same way that the equations are.

On the other hand, if the system of equations is overdetermined, then the system of equations may be restated as a nonlinear least squares problem to get a near solution to the system. It may also be desirable to weight the equations for the least squares problem to give greater emphasis to accurately satisfying the lower order moments. In the case where the number of points and the number of moments to be matched are not consistent, the error in the approximation to the objective function for the decision problem is more complex than otherwise.

5 Approximating Distributions of Known Form

The solutions to the system of moment equations are shown in Table 3 for the bivariate normal, and bivariate multinomial distributions. The number of points was chosen to be four and five. In the four point case, all moments through the third order (ten moments) can be satisfied. But with four points, there are twelve variables. Thus for each of the distributions, two cases are presented which employ alternative choices for two additional moments of order four to obtain a well determined system. In the first case, all ten moments of degree three or less, plus $<x^4>$ and $<y^4>$ are employed. In the second case, all moments of degree three or less, $<x^3y>$, and $<xy^3>$ are employed. In the five point case, the number of moments that can be satisfied is fifteen which happens to equal the number of moments of order four or less. As mentioned earlier, these equations are dependent. As a result, no solution exists to the full system for the bivariate normal case. To obtain a solution, we eliminate the moment $<x^2y^2>$ from the system of moment equation and solving the remaining system. Surprisingly, the moments for the multinomial case have nearly the same dependence relationship as the equation system. Thus, it was possible to compute a solution for which all of the moments are matched in their first four significant digits. The five-point approximations are displayed in Table 3 for both the joint normal and multinomial cases.

The systems of equations were solved using the MINOS nonlinear optimization system (Murtagh and Saunders). The problem was formulated as follows. The problem variables are the probabilities and associated values for the random variables in the discrete approximation. The objective was taken to be the sum of squares of the differences between the moments of the true distribution and the moments of the approximating distribution for all moments greater than the zero-th. The zero-th moment, requiring that the probabilities sum to one, is treated as a linear constraint. Otherwise, the problem is unconstrained. (To facilitate the solution process, bound constraints restricting probabilities to be nonnegative and less than one, and restricting the random variables to lie within the respective supports of the true distributions were also employed.)

Table 3: Multivariate Gaussian Quadrature Approximations

	Standard Normal			Multinomial		
	p	x	y	p	x	y
Four Points with $<x^4>$, $<y^4>$						
1	0.45412	0.74196	0.74196	0.39568	6.23816	6.23816
2	0.04588	-2.33441	2.33441	0.08026	1.30404	9.19697
3	0.45412	-0.74196	-0.74196	0.44380	3.80553	3.80553
4	0.04588	2.33441	-2.33441	0.08026	9.19697	1.30404
Four Points with $<x^3y>$, $<xy^3>$						
1	0.25000	1.00000	1.00000	0.31301	5.98328	6.72977
2	0.25000	1.00000	-1.00000	0.13355	1.98201	7.93674
3	0.25000	-1.00000	1.00000	0.39301	3.89396	3.56598
4	0.25000	-1.00000	-1.00000	0.16044	8.30273	2.69355
Five Points (excluding $<x^2y^2>$ for the normal)						
1	0.16667	1.73205	0.00000	0.20913	8.06222	4.51335
2	0.16667	-1.73205	0.00000	0.02896	0.00000	8.24387
3	0.05826	0.00000	2.32811	0.47493	3.84503	4.29489
4	0.36803	0.00000	0.61497	0.07583	7.07850	1.11506
5	0.24038	0.00000	-1.50581	0.21116	4.53111	8.04483

This approach is equally applicable to the approximation of either discrete or continuous distributions. Thus, when a true continuous distribution is unknown, but a sample is available, the procedure described here may be applied to produce an approximation to the sample using fewer points than the sample distribution. It is interesting that the discrete approximations do not necessarily share the symmetry properties of the original distributions. For instance, the first four-point normal distribution approximation is not symmetric with respect to reflections across either the axes, while the original distribution is invariant to such a reflection. Similarly, the second four-point approximation to the multinomial distribution is not symmetric about the forty five degree line where the two random variables are equal. Thus, the approximations are not unique.

6 Approximating Joint Normal Distributions

In cases where the distribution is multivariate joint normal and the system of nonlinear equations cannot conveniently be solved, an alternative approach to constructing a discrete approximation based on gaussian quadrature is available. The idea is as follows. Given a set of m jointly distributed normal random variables, determine a transformation which yields m independent normally distributed random variables which are linear functions of the original random variables. Use tabulated values (as in e.g., Stroud) or the approach of Miller and Rice to determine levels for the individual independent random variables. Invert the linear transformation to translate the levels for the independent random variables back to the original random variables. This results in an approximation to the original joint distribution that matches the same number of moments as are matched for the independent factors.

Let x be an m-dimensional vector of jointly distributed random variables with mean μ and variance Σ. (It is assumed that the matrix Σ is positive definite. If not, then techniques from principle component or factor analysis should be applied to reduce the set of random variables.) The discrete approximation may then be constructed by the following set of four steps.

1. Compute the Cholesky factors of $\Sigma = Q^t D Q$, where Q is orthonormal and D is diagonal.

2. Define the new random variables $z = Qx$ and note that the z's are independent and normally distributed with means equal to $Q\mu$ and variance covariance matrix D. (Letting q_i denote the i-th row of Q, $z_i = q_i x$ and the mean and variance of z_i are $q_i\mu$ and d_{ii}.

3. Use gaussian quadrature formulae to obtain discrete approximations for each of the z_i's. These discrete N-point approximations will be denoted $[\,{}^{j_i}p,\ {}^{j_i}z_i]_{j_i=1,\dots,N}$. The joint distribution of the z's is then denoted

$$\left\{\prod_{i=1}^{m} {}^{j_i}p, [\,{}^{j_1}z_1, {}^{j_2}z_2, \dots, {}^{j_m}z_m]^t\right\}_{j_1=1,\dots,N;\, j_2=1,\dots,N;\, \dots;\, j_m=1,\dots,N} = \{{}^{j}p, {}^{j}z\}_{j=1,2,\dots,N^m}.$$

4. Use the inverse transformation to obtain the final discrete distribution for the original set of jointly distributed random variables,

$$\{{}^{j}p, {}^{j}x\}_{j=1,2,\dots,N^m} = \{{}^{j}p, Q^t\,{}^{j}z\}_{j=1,2,\dots,N^m}.$$

The discrete distribution constructed by this approach will have moments of order zero through $2N-1$ which will match those of a joint normal distribution with mean μ and variance-covariance matrix Σ.

Table 4: Normal Distribution Approximated by Direct Solution of Moment Equations (Seven Points) and by Decomposition (Nine Points)

Point Number	Direct Equation Solution Approach			Decomposition Approach		
	p	x	y	p	x	y
1	0.08333	-2.82643	0.07520	0.11111	1.73205	0.00000
2	0.08333	2.82643	-0.07520	0.02778	1.73205	2.44949
3	0.08333	1.50531	1.69232	0.11111	0.00000	2.44949
4	0.08333	-1.32112	1.76842	0.02778	-1.73205	2.44949
5	0.50000	0.00000	0.00000	0.11111	-1.73205	0.00000
6	0.08333	1.32112	-1.76842	0.02778	-1.73205	-2.44949
7	0.08333	-1.50531	-1.69323	0.11111	0.00000	-2.44949
8				0.02778	1.73205	-2.44949
9				0.44444	0.00000	0.00000

The results of applying this procedure as well as the procedure described in the previous section are displayed in Table 4 for a bivariate normal distribution with correlated random variables. The distributions for the transformed, independent normal random variables are approximated by three points each. Hence, the discrete distribution based on this decomposition approach employs nine points and has all moments up to the fifth order correctly specified. The alternative method based on direct solution of the nonlinear moment equations employs seven points and also exactly satisfies all moments up to the fifth order. In this bivariate case the difference in the number of points employed by these two approaches is small. However, the difference can, in general, be quite large since for a fixed number of moments, the number of points required increases exponentially in the number of random variables. For instance, consider the case where there are five random variables and the goal is to have all moments of order three or less exactly correct. In this case, twelve points are required by the method based on direct solution of the nonlinear equations, but 32 points are needed for the approach based on decomposition.[3] Thus, even though the numerical effort required for direct solution of the nonlinear equations is much higher than for decomposition, the former approach may still be desirable if the cost of having a distribution with more points in the decision problem is substantial.

7 Conclusions

The accuracy of the solution of a stochastic decision problem depends critically on how accurately the probability distribution is expressed in the solution process. When practical limits to computation make using the true distribution impractical, the distribution must be approximated by a relatively small number of points with positive mass. Conventional approaches to making such an approximation may grossly understate variances and other higher order moments.

[3] This comparison is not quite fair. With the direct solution approach, all moments up to the given order are satisfied exactly. With the decomposition approach, all moments involving powers of the random variables up to the given order are satisfied exactly. Thus, in our example in Table 4, $\langle x^5y^5 \rangle$ is correct for the decomposition approach, but not for the direct solution approach.

One approach to constructing practical discrete approximations to distributions is to choose the points and probabilities so as to match as many lower order moments of the original distribution as needed. When there is a single random variable or when random variables are independent, this is equivalent to the gaussian quadrature approach to numerical integration. The method is generalized here to the case of jointly distributed random variables. This method can be applied to any distribution for which the low order moments through the degree desired can be determined, regardless of whether the original distribution is discrete or continuous. An alternative method which produces approximate distributions with more discrete points, but which is easier to compute, is also available for the case where the random variables are distributed jointly normal.

References

1. Miller, A.C. and T.R. Rice, "Discrete Approximations of Probability Distributions," *Management Science*, 29(3):352-362.

2. Murtagh, B.A. and M.A. Saunders, "MINOS 5.1 User's Guide," Systems Optimization Laboratory Technical Report #83-20R, Department of Operations Research, Stanford University, January 1987.

3. Stroud, A.H. and D. Secrest, *Gaussian Quadrature Formulas*, Prentice Hall, Englewood Cliffs, N.J., 1966.