

A Comparison of Methods for the Treatment of Uncertainties in the Modeling of Complex Systems

by

Isaac Trefz

B.S., Swarthmore College (1994)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1998

© Massachusetts Institute of Technology, MCMXCVIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic
copies of this thesis document in whole or in part, and to grant others the right to do so.

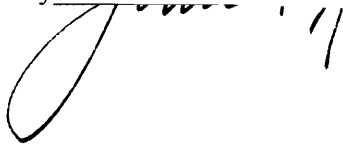
Author _____



Department of Electrical Engineering and Computer Science

May 26, 1998

Certified by _____



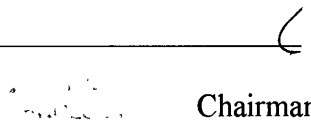
J. G. Kassakian
Professor of Electrical Engineering
Primary Thesis Supervisor

Certified by _____



T. M. Jahns
Senior Lecturer of Electrical Engineering
Secondary Thesis Supervisor

Accepted by _____


Chairman, Departmental Committee on Graduate Students

Arthur C. Smith

JUL 23 1998

LIBRARIES

A Comparison of Methods for the Treatment of Uncertainties in the Modeling of Complex Systems

by

Isaac Trefz

Submitted to the Department of Electrical Engineering and Computer Science
on May 26, 1998 in partial fulfillment of
the requirements for the degree of
Master of Science in Electrical Engineering

Abstract

Engineers working on complicated system integration problems in industries such as automotive or aerospace often face a need to understand the impact of uncertainties in their component models, particularly during the early stages of development projects. If such uncertainties are modeled as continuous random variables, system model calculations create the need to evaluate the probability characteristics of functions (i.e., sums, products, etc.) of these random variables. Evaluation of analytical expressions for the resulting probability density functions is often intractable, and even direct numerical evaluations can require excessive computation times using modern computers. As an alternative, there are approximation techniques which hold the prospect of significant reductions in the execution times in return for modest reductions in the accuracy of the resulting density functions.

Methods for approximating mathematical functions of independent random variables are explored based on the assumption that the operand random variables can be modeled as beta distributions. Attention is focused on the arithmetic operations of multiplication, division, and addition of two independent random variables, and the exponentiation of one random variable to a scalar power. Execution times for these approximate evaluation techniques are compared to those using direct numerical calculation and, alternatively, to those using a Monte Carlo approach. Comparisons are made of the accuracies of these alternative approximation techniques to the results of direct numerical calculations for a wide range of beta distribution parameters. These tests indicate that the approximation techniques generally provide good accuracy for addition, multiplication, and exponentiation operations, but the approximation accuracy is poorer for the quotient operation. This limitation is not expected to pose a major impediment to the useful application of these approximations in system model calculations.

Primary Thesis Supervisor: J. G. Kassakian
Title: Professor of Electrical Engineering

Secondary Thesis Supervisor: T. M. Jahns
Title: Senior Lecturer of Electrical Engineering

Dedication

to

my Grandparents

and

my Parents

Acknowledgments

First and foremost I would like to thank Thomas Neff of Daimler-Benz Forschung in Frankfurt a.M. for developing the concept for this work [see 12] and for his supervision and assistance. I hope he finds my work useful when writing his Ph.D. thesis. On the same note, thanks are due to Dr. Michael Kokes and Dr. Hans-Peter Schöner of Daimler-Benz for finding the money to pay my way here at MIT. This research would not have been possible without the financial support of Daimler-Benz AG.

Next in line for thanks is Dr. Thomas Jahns. Working closely with a company on the other side of the ocean had its snares, but he managed to work them all out. He also took the time to offer advice whenever I needed it and carefully read several drafts of this thesis.

Professor John Kassakian wasn't afraid to take a chance on me from the beginning. He also let me TA his Power Electronics course and gave me two paid vacations, one to Europe and one to Phoenix. Thanks for all of the good times, John.

Acknowledgments to Khurram Khan Afridi for writing MAESTrO and helping me get my foot in the door with Daimler-Benz. Khurram was basically my mentor here at MIT, and everything I know about wheeling and dealing I learned from him.

Big Thanks to Ankur Garg. MIT is a rough place, and I doubt that I would have survived 6.241, 6.341 and the PWE without her. Her help with this thesis has also proved invaluable. Most importantly though, she is simply a good friend and all-around class act.

The basement of Building 10 may be a hole but Vahe 'Iron Man' Caliskan and James Hockenberry made it a bearable one. Vahe knows everything there is to know about power electronics and is the nicest, most considerate person I've ever met. James and I spent hours on end reminiscing about our alma mater, Swarthmore College, and discussing society's ills. Thanks guys.

Our secretaries here in LEES, Vivian Mizuno, Karin Janson-Strasswimmer, and Sara Wolfson, also deserve thanks for their moral support and willingness to help out with anything and everything.

Contents

1 Introduction	15
1.1 System Analysis Using MAESTrO	16
1.2 Objectives and Technical Approach	18
1.3 Organization of this Thesis	19
2 Background	21
2.1 Basic Probability Theory	21
2.1.1 Expression of Uncertainty as a Random Variable.....	21
2.1.2 Moments and Central Moments of a Continuous Random Variable	22
2.1.3 Characteristic Function of a Continuous Random Variable.....	23
2.2 The Generalized Beta Distribution.....	25
3 Analytical Approximation Methods	31
3.1 Springer's Methods	32
3.1.1 Linear Combinations of Two Independent Random Variables	32
3.1.2 Multiplication of Two Independent Random Variables.....	34
3.1.3 Exponentiation of One Random Variable	36
3.1.4 Division and Subtraction of Two Random Variables	36
3.2 McNichols' Methods.....	37

4 Numerical Methods	43
4.1 Analytical Solutions to Functions of Random Variables	43
4.1.1 Analytical Convolution of Two Independent Random Variables	43
4.1.2 Analytical Exponentiation of One Random Variable.....	46
4.2 Direct Numerical Approximation to Functions of Random Variables	46
4.2.1 Numerical Convolution of Two Independent Random Variables.....	46
4.2.2 Numerical Exponentiation of One Random Variable	49
4.3 Monte Carlo Approximation of Functions of Random Variables.....	49
5 Implementation and Experimental Results	51
5.1 Calculation of a Beta Density Function from its Beta Parameters.....	51
5.2 Chi-Squared and Least-Squares Tests.....	54
5.3 Implementation and Comparison of Numerical Methods	56
5.3.1 Numerical Convolution of Two Independent Random Variables.....	57
5.3.2 Numerical Exponentiation of One Random Variable	59
5.3.3 Implementation and Exploration of Monte Carlo Methods	59
5.4 Implementation and Comparison of Analytical Approximation Methods.....	60
5.4.1 Calculation of Beta Parameters from Central Moments.....	61
5.4.2 Implementation of Springer’s Methods.....	61
5.4.3 Implementation of McNichols’ Methods	62
5.4.4 Description of Test Batteries	62
5.4.5 Experimental Comparison of Springer’s and McNichols’ Methods.....	64
5.5 Comparison of Execution Times.....	71
6 Conclusions and Further Research	73

List of Figures

Figure 1.1: A simple electrical system that one may analyze using MAESTrO is shown above. This system is comprised of of an alternator, two motor loads, and their associated wiring.	16
Figure 1.2: Probability density functions that may be used to represent the cost of the alternator (left), motor1 (center), and motor2 (right) of the system shown in Figure 1.1 are shown above.....	17
Figure 1.3: A probability density function representing uncertainty in the cost of the system of Figure 1.1 is shown above.....	17
Figure 2.1: The beta density function plotted on the interval [0,1] for various values of α and β both greater than 1 is shown above.	26
Figure 2.2: The beta density function plotted on the interval [0.1] for various values of α and β where either $\alpha < 1$ or $\beta < 1$ (left) or both $\alpha < 1$ and $\beta < 1$ (right) is shown above.	27
Figure 3.1: The general method used to approximate a function of two independent beta variables is represented here. The step inside the circle will differ depending on the functional form of the function to be approximated and the approximation method used (i.e. Springer's or McNichols').	31
Figure 5.1: Plots of $f_b = (x; 2 \times 10^4, 2 \times 10^4, 100, 200)$ (left) and $f_b = (x; 75, 75, 100, 200)$ (right) are shown above. Note how much sharper the density function on the left is compared to the density function on the right.....	53
Figure 5.2: A flowgraph of the chi-squared routine used to test each all approximation routines is shown above.	55
Figure 5.3: A flowgraph of the least-squares routine used to test each of Springer's and McNichols' approximation routines is shown above.....	57
Figure 5.4: Plots of the product of $f_b(x; 2, 1, 0, 1)$ and $f_b(x; 3, 1, 0, 1)$ generated using numerical convolution (left) and this same curve superimposed upon $f_b(x; 2, 2, 0, 1)$ (right) are shown above. The plot on the left shows that the technique yields smooth curves, while the plot on the right shows that accurate curves are yielded through this technique as well.	58

- Figure 5.5: The product of $f_1 = f_b(x;2,1,0,1)$ and $f_2 = f_b(x;3,1,0,1)$, calculated using the Monte Carlo method (solid), and the known product $g = f_b(x;2,2,0,1)$ (dashed) are plotted above. Comparison with Figure 5.4 shows that the convolution based methods described in the previous section are far superior in terms of the smoothness of the result than the Monte Carlo method described in this section. 60
- Figure 5.6: Plots of the approximation to the product of $f_b(x;4.1,8.15,0,1)$ and $f_b(x;9,9.33,0,1)$ shown on the interval $[0,0.6]$ are shown above for Springer's method (left) and McNichols second order method (right). It can be seen that Springer's method yields the better approximation. 67
- Figure 5.7: Plots of Springer's approximation to the product of $f_b(x;1.24,3.81,0,1)$ and $f_b(x;1.12,4.46,0,1)$ and the numerically generated benchmark density function on the interval $[0,0.1]$ are shown above. In this case, Springer's method yields a result of $f_b(x;0.57,2.170,0.00,1.79)$ and a least-squares metric of 2.06. 68
- Figure 5.8: Plots of Springer's approximation to the sum of $f_b(x;9.25,2.11,0,1)$ and $f_b(x;1.12,4.33,0,1)$ and the numerically generated benchmark density function on the interval $[0,2]$ are shown above. In this case, Springer's method yields a result of $f_b(x;49.75,49.75,-0.92,2.96)$ and a least-squares metric of 0.0177. 68
- Figure 5.9: Plots of the approximation to the quotient of $f_b(x;4.74,3.75,1,2)$ and $f_b(x;8.87,1.36,1,2)$ on the interval $[0.4,1.5]$ are shown above for the three methods tested. Both McNichols' first and second order approximations are more accurate than Springer's, but it is difficult to discern that from this figure..... 69
- Figure 5.10: Plots of Springer's approximation to the quotient of $f_b(x;8.15,1.53,1,2)$ and $f_b(x;6.43,1.45,1,2)$ and the numerically generated benchmark density function on the interval $[0.5,1.5]$ are shown above. In this case, Springer's method yields a result of $f_b(x;6.10,6.10,0.65,1.39)$ and a least-squares value of 0.2151. 70

List of Tables

Table 1.1: Assumptions about the minimum, maximum, and most likely costs of the alternator and motors of the system shown in Figure 1.1 are given above.....	17
Table 3.1: The cumulants of $Y = \varphi(X_1, X_2, \dots, X_n)$ as a function of the central moments of X_1, X_2, \dots, X_n for McNichols' first order approximation (see Equation (3.22))	38
Table 3.2: The cumulants of $Y = \varphi(X_1, X_2, \dots, X_n)$ as a function of the central moments of X_1, X_2, \dots, X_n for McNichols' second order approximation (see Equation (3.23)).....	39
Table 3.3: The cumulants of $Y = X_1 X_2$ as a function of the central moments of X_1 and X_2 for McNichols' first order approximation.....	40
Table 3.4: The cumulants of $Y = X_1 X_2$ as a function of the central moments of X_1 and X_2 for McNichols' second order approximation	40
Table 3.5: The cumulants of $Y = X_1/X_2$ as a function of the central moments of X_1 and X_2 for McNichols' first order approximation.....	41
Table 3.6: The cumulants of $Y = X_1/X_2$ as a function of the central moments of X_1 and X_2 for McNichols' second order approximation	41
Table 3.7: The cumulants of $Y = X^m$ as a function of the central moments of X_1 for McNichols' first order approximation.....	42
Table 3.8: The cumulants of $Y = X^m$ as a function of the central moments of X_1 for McNichols' second order approximation	42
Table 5.1: The beta parameters and number of tests run for the first battery of random tests on Springer's and McNichols' approximation algorithms are shown above. All of the beta shape parameters, α and β , were produced using a random number generator...	63
Table 5.2: The beta parameters and number of tests run for the second battery of random tests on Springer's and McNichols' approximation algorithms are shown above. All of the beta interval boundaries, a and b , were produced using a random number generator. .	64
Table 5.3: The results of the first battery of tests run on Springer's and McNichols' approximation methods are shown above (see Table 5.1). For each arithmetic operation and approximation method, the number of trials out of one hundred that failed a chi-squared test is given, along with the average over all one hundred trials of	

the least-squares test (see Equation (5.11)). Note that McNichols' first and second order methods are identical to Springer's method for the case of addition (see Section 3.2)..... 65

Table 5.4: The results of the second battery of tests run on Springer's and McNichols' approximation methods are shown above (see Table 5.2). For each arithmetic operation and approximation method, the number of trials out of one hundred that failed a chi-squared test is given, along with the average over all one hundred trials of the least-squares test (see Equation (5.11)). Note that McNichols' first and second order methods are identical to Springer's method for the case of addition (see Section 3.2)..... 66

Table 5.5: Execution times in seconds of all of the numerical and analytical methods explored for generating approximations to products, quotients, sums, and powers of independent beta random variables are shown above. 71

1 Introduction

Enterprises that perform system integration must be able to evaluate system components in early development stages, when more than half of total product costs is determined [3]. Manufacturers of automobiles and trucks are excellent examples of such enterprises. At present almost all development and determination of subsystem costs is done by the suppliers of individual subsystems, but there is a need for the producer of the overall system to plan and coordinate the development of these subsystems. The system integrator must therefore be able to effectively model economic aspects, such as life-cycle costs, and technical aspects, such as electrical losses, of system components. During the early development phase when system alternatives are discussed, not all of the information required for an effective comparison of various system topologies is available. Many of the attributes of various system components can only be represented by estimated values. Nevertheless, important product decisions are made on the basis of these inexact estimates. It is therefore desirable to render visible the uncertainty apparent in all available data and the results of any system analysis before major decisions are made.

One method of modeling uncertainties in component attributes is to use continuous random, or stochastic, variables. Uncertainty in overall system attributes may then be calculated and also represented in the form of continuous random variables. The representation and calculation of system attributes using random variables opens up the possibility of assessing the risks involved with a given project. There are presently several well known mathematical methods for calculating functions of several random variables [see 16]. Unfortunately, many of these methods require long computation times or are not well suited for implementation on a computer [see 2 for an example of this]. When the systems to be modeled are large and results are needed quickly, long computation times and large data structures become a major hindrance. The purpose of this thesis is to investigate methods for efficiently, accurately, and quickly carrying out mathematical operations with random variables, which will lend themselves well to computer implementation and integration into software.

1.1 System Analysis Using MAESTrO

MAESTrO [19] is a software tool specifically designed to analyze alternative electrical system configurations for automobiles and trucks. Specifically, MAESTrO is capable of calculating estimates of the total cost, weight, reliability, and efficiency of an entire electrical system as well as of its generation, storage, distribution, and utilization subsystems. Figure 1.1 shows a simple electrical system that one may analyze using MAESTrO. This system is comprised of an alternator, two motor loads, and their associated wiring. Suppose we wish to calculate the cost of the entire system by adopting the simple cost model assumption that the cost of the total system is equal to the sum of the costs of all six system components. At present, the cost of each component, and thus of the entire system, can only be calculated deterministically in MAESTrO. All of the input parameters describing each component to MAESTrO are scalars, and the cost of the system is therefore output by the software as a scalar. However, this scalar result cannot meaningfully represent the uncertainty present in the cost of the system due to uncertainty present in the parameters used to calculate the costs of each individual component.

As an example, we will now proceed to probabilistically calculate the cost of the system shown below. Suppose that each of the three wires is known to cost \$5, and we make the assumptions shown in Table 1.1 about the minimum possible, maximum possible and most likely costs of the other three components.

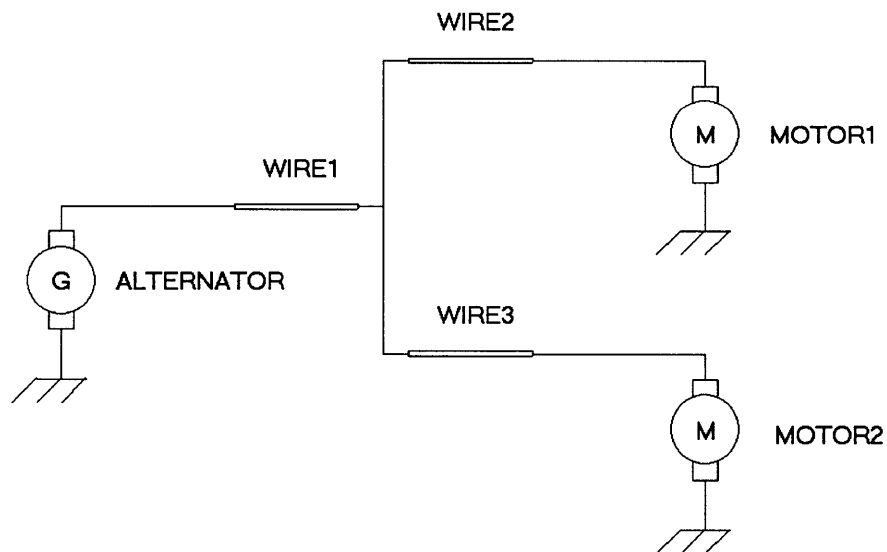


Figure 1.1: A simple electrical system that one may analyze using MAESTrO is shown above. This system is comprised of an alternator, two motor loads, and their associated wiring.

Component	Minimum Cost	Maximum Cost	Most Likely Cost
Alternator	\$30	\$40	\$35
Motor1	\$20	\$30	\$27
Motor2	\$5	\$9	\$7

Table 1.1: Assumptions about the minimum, maximum, and most likely costs of the alternator and motors of the system shown in Figure 1.1 are given above.

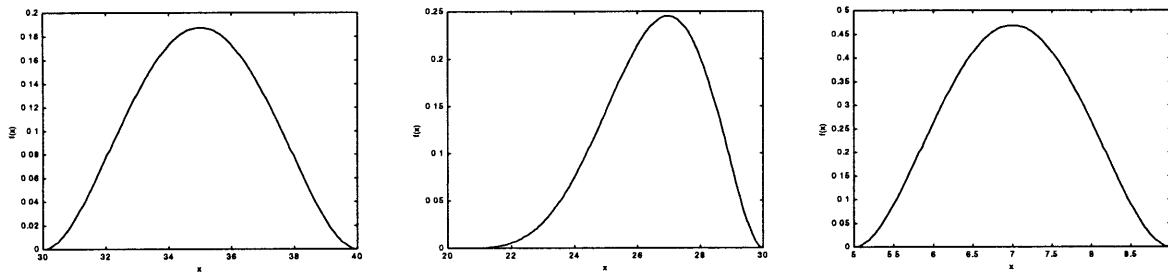


Figure 1.2: Probability density functions that may be used to represent the cost of the alternator (left), motor1 (center), and motor2 (right) of the system shown in Figure 1.1 are shown above.

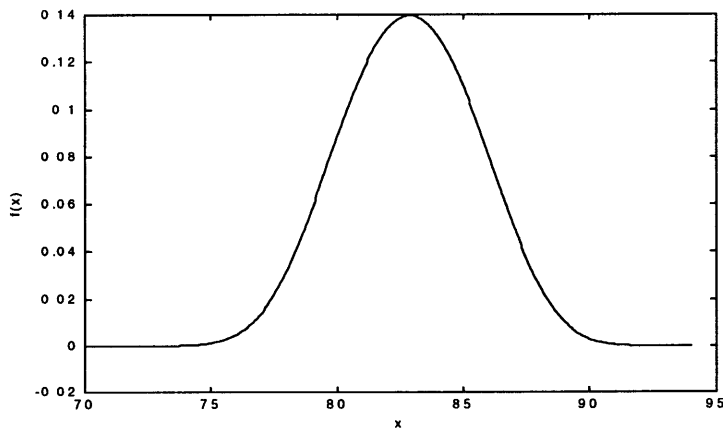


Figure 1.3: A probability density function representing uncertainty in the cost of the system of Figure 1.1 is shown above.

The cost assumptions shown in Table 1.1 can be approximated by the probability density functions shown in Figure 1.2. We can use convolution to add these three density functions to each other and to the total cost of the three wires (\$15), to obtain a density function representing the cost of the entire system. This result is shown in Figure 1.3 and yields a great deal of valuable information about the uncertainty inherent in the cost of the entire system. For example,

it is clear that this cost will not be less than \$75, nor exceed \$91. The expected value of the cost can also be calculated to be \$82.85.

We can also find the probability that the cost of the system will be between two values, x_1 and x_2 by integrating the above distribution over the interval $[x_1, x_2]$. For example, suppose we have defined a maximum cost target for our system of \$85. We can integrate the above distribution from $x = \$0$ to $x = \$85$ in order to find the probability (77.5%) that we will meet this cost target. Knowing that a system design may meet a desired cost target with a probability of only 77.5% long before the system is about to enter production could prove to be a valuable piece of information at early design stages. Unfortunately, such information cannot be obtained with the current implementation of MAESTrO.

The goal of this thesis is to explore means of extending MAESTrO to allow users to define component parameters as probability density functions, as opposed to simple scalars. From these density functions representing component parameters, we wish to compute probability density functions representing subsystem and system attributes such as cost, weight, reliability, and efficiency. In order to be able to do so, we must find fast and effective methods to compute mathematical functions of random variables.

1.2 Objectives and Technical Approach

The overall objective of this thesis is to investigate methods of approximating probability densities of functions of independent random variables for the purpose of using in the MAESTrO software tool. Stated mathematically, if X_1, X_2, \dots, X_n are independent real random variables with density functions $f_1(x), f_2(x), \dots, f_n(x)$ and $Y = \varphi(X_1, X_2, \dots, X_n)$, where φ is any arbitrary real function, we would like to find a reasonable approximation to the probability density function of the random variable Y , $g(y)$, using the density functions of X_1, X_2, \dots, X_n .

While it is theoretically possible in many instances to find a closed-form analytical expression for $g(y)$, this expression can be very complicated, difficult to derive, and impossible to evaluate on a computer. It is therefore desirable to find an acceptably good approximation to $g(y)$ using simple numerical algorithms that would require a fraction of the computational time and resources of their analytical counterpart. This thesis will compare methods for approximating the density function, g , of the function of independent random variables, φ .

The technical approach adopted in this work is to investigate the following three methods for performing calculations with random variables and to compare them with regard to their relative speed and accuracy:

1. Direct numerical calculation of $g(y)$ based on analytical expressions
2. Monte Carlo approximation to $g(y)$
3. Analytical approximation to $g(y)$ using the moments or cumulants of X_1, X_2, \dots, X_n

The mathematical operations that will be investigated are products, quotients, and linear combinations of two independent random variables, and exponentiation of one random variable to a scalar power. While analytic expressions for all of these operations exist, they typically involve convolution integrals that are difficult or impossible to evaluate analytically. These expressions may however be used as a basis to directly generate numerical approximations to $g(y)$.

A Monte Carlo approach can serve as an alternative to direct numerical approximation of analytic expressions and has the advantage that it does not require there to be an analytic expression available for the density function being approximated. This approach is considerably more complicated than direct numerical calculation but can in some cases lead to improvements in speed and accuracy.

Although direct numerical calculation and Monte Carlo calculation of functions of independent random variables may require less time and computational resources than a direct analytical approach, the computational savings that these methods offer may still not be enough if we have models that require many stochastic calculations or require storage of many stochastic variables. Fortunately, methods are available to approximate the density function of $Y = \varphi(X_1, X_2, \dots, X_n)$ using simple calculations and the moments of X_1, X_2, \dots, X_n . Three of these methods, one developed by M. Springer [16] and two developed by G. McNichols [11] can offer significant savings in computation time and resources over direct numerical calculation or a Monte Carlo approach.

1.3 Organization of this Thesis

As mentioned in the previous section, this thesis will explore three methods for approximating the density function, $g(y)$, of the random variable $Y = \varphi(X_1, X_2, \dots, X_n)$, where X_1, X_2, \dots, X_n are independent random variables and φ is a real function. Chapter 2 begins with an introduction to

the basic concepts of probability theory that are necessary to the derivation of the analytical approximation methods of Springer and McNichols, and concludes with a description of the generalized beta distribution which is used for all calculations in this thesis.

Chapter 3 develops the analytical approximation methods that will be used to approximate the density function of Y , $g(y)$. The general method for approximating a function of beta distributions with a beta distribution is introduced. Subsequently, Springer's and McNichols' methods, which are used to calculate the moments of the approximating beta distribution, are developed out of the theory of Chapter 2.

Chapters 4 develops the numerical and Monte Carlo methods that will be used to approximate the density function of Y , $g(y)$. Analytical expressions for the density functions of products, quotients, sums, and differences of two independent random variables, as well as exponentiation of one random variable to a scalar power, are derived, and algorithms are developed to generate numerical approximations based on these analytical expressions.

Chapter 5 begins with a discussion of implementation issues surrounding the direct numerical and Monte Carlo algorithms of Chapter 4 and compares the two approximation methods on the basis of the smoothness and accuracy of the approximations they generate. Implementation issues surrounding the analytical approximation methods of Chapter 3 are then discussed. Two batteries of tests are used to compare Springer's and McNichols' methods, and their results are presented. The chapter concludes with a comparison of the execution times of all the approximation methods under study.

Chapter 6 summarizes the results of Chapter 5 in the context of the extension of MAESTrO to perform calculations with random variables. The conclusion is reached that while direct numerical techniques are still too slow for software implementation, Springer's approximation techniques hold promise for the addition, multiplication, and exponentiation operations. The approximation accuracy for the quotient operation is found to be poorer. However, this limitation is not expected to pose a major impediment to the useful application of these approximations in system model calculations.

2 Background

Basic probability theory and the beta distribution are fundamental to the derivation of the analytic and numeric methods that are the subject of exploration of this thesis. Section 2.1 introduces the concepts of a random variable, the moments and central moments of a random variable, and the characteristic function of a random variable. These basic ideas are then applied to generalized beta random variables in Section 2.2 in order to show that a beta random variable can be completely and uniquely characterized by its first four moments.

2.1 Basic Probability Theory

This section introduces basic concepts of probability theory necessary to develop the numerical approximation methods described in Chapter 3. The section begins with the definition of a random variable and its associated distribution and probability density functions. Expressions for calculating the moments of a random variable from its density function and its first four central moments from its moments are then presented. Finally, it is shown through the use of the characteristic function that a reasonable approximation to the density function of a random variable may be obtained from the random variable's moments.

2.1.1 Expression of Uncertainty as a Random Variable

An uncertain quantity can be modeled as a random variable. Probability theory differentiates between discrete and continuous random variables. A discrete random variable may only assume a finite or countable number of values, whereas a continuous random variable may assume an infinite number of values within the interval upon which it is defined [5]. Random variables shall be denoted henceforth by uppercase letters and the values that they may assume by lowercase letters.

Every random variable has a non-negative, monotonically non-decreasing distribution function, $F(x)$, associated with it defined as:

$$F(x) = P(X \leq x) \tag{2.1}$$

where $F(-\infty) = 0$ and $F(\infty) = 1$. If $F(x)$ is continuous and everywhere differentiable, we may define the probability density function of a random variable, $f(x) = dF(x)/dx$. A probability density function has the following properties:

$$f(x) \geq 0 \quad (2.2a)$$

$$P(x_1 \leq X < x_2) = \int_{x_1}^{x_2} f(x) dx \quad (2.2b)$$

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (2.2c)$$

2.1.2 Moments and Central Moments of a Continuous Random Variable

The r^{th} moment of a random variable about zero is given by [9]:

$$\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx \quad (2.3)$$

The first moment about zero of a random variable, μ'_1 , is called its mean or expected value.

We may also define the set of moments about the mean of a random variable and call these the central moments. The r^{th} central moment of a random variable is therefore given by:

$$\mu_r = E((X - \mu'_1)^r) = \int_{-\infty}^{\infty} (x - \mu'_1)^r f(x) dx \quad (2.4)$$

The following is true of the first and second central moments:

$$\mu_1 = E(X - \mu'_1) = E(X) - \mu'_1 = 0 \quad (2.5a)$$

$$\mu_2 = E(X - \mu'_1)^2 = E(X^2) - \mu_1'^2 = \mu_2' - \mu_1'^2 \quad (2.5b)$$

The second central moment of a random variable is also known as the variance and is often written as σ^2 .

Similarly, the third and fourth central moments of a random variable may be expressed in terms of the moments:

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 \quad (2.6a)$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 \quad (2.6b)$$

2.1.3 Characteristic Function of a Continuous Random Variable

The expectation of the random variable e^{itX} considered as a function of the real variable t is known as the characteristic function of X [13]. The characteristic function of a random variable X has many important properties and is crucial in establishing a connection between the moments of a distribution and the distribution itself. A random variable's distribution function is completely determined by its characteristic function, and the characteristic function of a random variable may be expressed as a function of the moments. Thus, the distribution function of a random variable may be completely specified through moments. As this fact will form the basis for all of the analytic approximation methods developed in Chapter 3, a further exploration of the characteristic function is necessary.

The characteristic function of a real random variable X is given by [13]:

$$\phi(t) = E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx, \text{ where } i = \sqrt{-1} \quad (2.7)$$

Since $|e^{itx}| = 1$ for all real t , the above integral will converge for any real random variable. The density function of X may therefore be expressed as the following Fourier integral:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t) dt \quad (2.8)$$

Given that the characteristic function is differentiable, its r^{th} derivative is given by:

$$\frac{d^r \phi}{dt^r} = i^r \int_{-\infty}^{\infty} x^r e^{itx} f(x) dx \quad (2.9)$$

By letting $t = 0$ and comparing with Equation (2.3), we obtain:

$$\mu'_r = i^{-r} \left. \frac{d^r \phi}{dt^r} \right|_{t=0} \quad (2.10)$$

Given that a random variable has moments up to order n , we may express its characteristic function as a Maclaurin series:

$$\phi(t) = 1 + \sum_{r=1}^n \frac{i^r \mu'_r}{r!} t^r + R_n(t) \quad (2.11)$$

However, as $n \rightarrow \infty$, the residual term, R_n , converges to 0:

$$\lim_{n \rightarrow \infty} R_n(t) = \lim_{n \rightarrow \infty} \frac{\mu'_{n+1}}{(n+1)!} t^{n+1} = 0 \quad (2.12)$$

Thus, the characteristic function may be expressed as an infinite series of the moments:

$$\phi(t) = \lim_{n \rightarrow \infty} \left\{ 1 + \sum_{r=1}^n \frac{i^r \mu'_r}{r!} t^r \right\} \quad (2.13)$$

We have shown that if an infinite number of the moments of a random variable exist and are known, then its characteristic function, and thereby its distribution and density functions, may be completely determined. However, in almost all practical cases not all of the moments of a given random variable are available. One can however approximate the density function of a random variable by using a finite number of moments.

Suppose we wish to approximate the density function of a random variable using the following finite power series [9]:

$$f \cong \sum_{j=0}^n a_j x^j \quad (2.14)$$

If f is defined on the closed interval $[a, b]$, the coefficients of a_r may be determined by the method of least-squares. In order to do so, the following expression must be minimized:

$$g = \int_a^b \left(f - \sum_{j=0}^n a_j x^j \right)^2 dx \quad (2.15)$$

Differentiating with respect to the coefficient a_r , where r is any integer on the interval $[0, n]$, and setting the result to 0 yields:

$$-2 \int_a^b \left(f - \sum_{j=0}^n a_j x^j \right) x^r dx = 0, \quad r = 0, \dots, n \quad (2.16a)$$

$$\Leftrightarrow \int_a^b x^r f dx = \mu'_r = \int_a^b \sum_{j=0}^n a_j x^{j+r} dx, \quad r = 0, \dots, n \quad (2.16b)$$

If two distributions have equal moments up to an arbitrary order n , then they will have the same least-squares approximation, as the coefficients a_j are determined by the moments. This result opens up the possibility of approximating a function of one or more random variables by its moments. According to Kendall and Stuart, the first four moments of a distribution function are sufficient to yield a reasonably good approximation to that distribution function, given that we have some information about the type of the resulting distribution function. In the following section we will describe a distribution function that is uniquely specified by its first four moments.

2.2 The Generalized Beta Distribution

The generalized beta distribution [15] is able to represent the sort of uncertainties that arise in practical industrial applications quite well, since it is defined over a finite interval (as opposed to the Gaussian distribution) and provides a close approximation to a large number of other types of distribution functions.

The generalized beta distribution is defined uniquely by four parameters, α , β , a , and b , and its density function is given by the following relationship:

$$f_B(x; \alpha, \beta, a, b) = \frac{1}{(b-a)^{\alpha+\beta-1}} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (x-a)^{\alpha-1} (b-x)^{\beta-1} \quad (2.17)$$

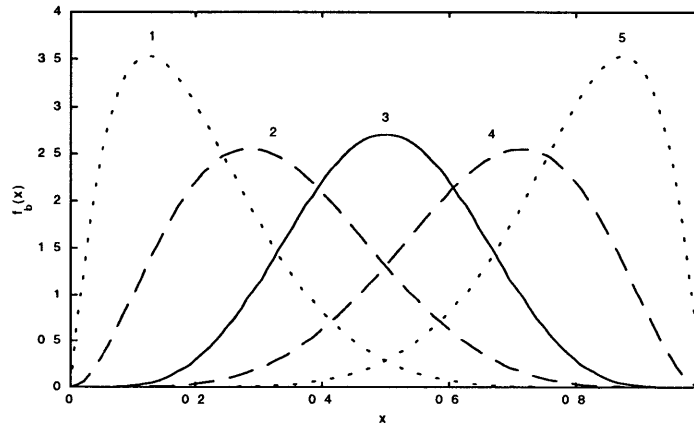
where $a \leq x \leq b$ and $\alpha, \beta > 0$. The Γ -function is defined as:

$$\Gamma(x) = (x-1)! = \int_0^{\infty} e^{-t} t^{x-1} dt \quad (2.18)$$

The beta distribution is useful for representing a wide range of distribution functions over the finite interval $a \leq x \leq b$. The parameters α and β are both shape parameters. When α and β are both greater than 1 (see Figure 2.1), the distribution will have only one peak at:

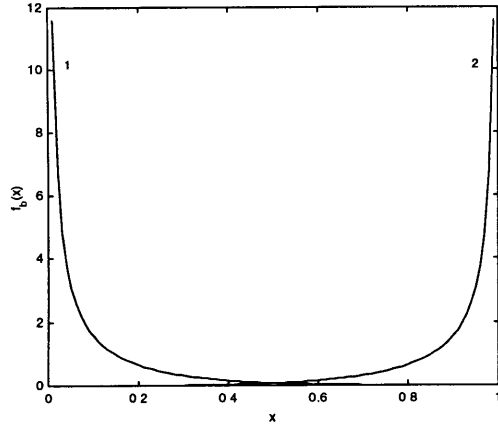
$$x_{peak} = \frac{a(\beta-1) + b(\alpha-1)}{\alpha + \beta - 2}, \text{ for } \alpha > 1 \text{ and } \beta > 1 \quad (2.19)$$

When α and β are both less than 1, the function assumes a U-shape. When $\alpha \geq 1$ and $\beta < 1$ it is J-shaped, and when $\beta \geq 1$ and $\alpha < 1$ the function takes the shape of an inverse J (see Figure 2.2). The function becomes the uniform distribution when $\alpha = \beta = 1$.

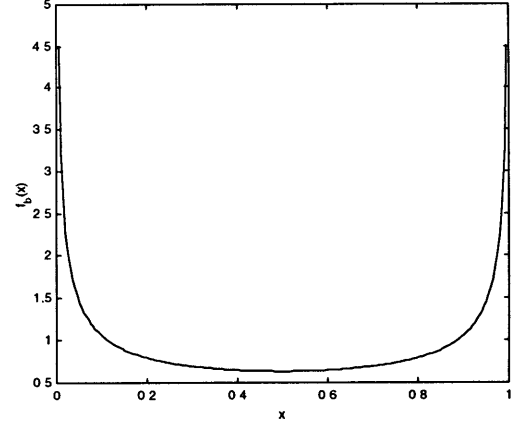


- | | | | |
|---|------------------|---|------------------|
| 1 | $f_b(x;2,8,0,1)$ | 4 | $f_b(x;6,3,0,1)$ |
| 2 | $f_b(x;3,6,0,1)$ | 5 | $f_b(x;8,2,0,1)$ |
| 3 | $f_b(x;6,6,0,1)$ | | |

Figure 2.1: The beta density function plotted on the interval $[0,1]$ for various values of α and β both greater than 1 is shown above.



- 1 $f_b(x;0.25,4,0,1)$
 2 $f_b(x;4,0.25,0,1)$



$f_b(x;0.5,0.5,0,1)$

Figure 2.2: The beta density function plotted on the interval [0,1] for various values of α and β where either $\alpha < 1$ or $\beta < 1$ (left) or both $\alpha < 1$ and $\beta < 1$ (right) is shown above.

In order to be able to approximate a distribution whose first four moments are known with a beta distribution, it is necessary to establish a relationship between the first four moments of a beta distribution and its four beta parameters. McNichols [11] has calculated the mean, second, third, and fourth central moments of the generalized beta distribution as a function of the beta parameters:

$$\mu'_1 = a + (b - a) \left[\frac{\alpha}{\alpha + \beta} \right] \quad (2.20a)$$

$$\mu_2 = \sigma^2 = (b - a)^2 \left[\frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} \right] \quad (2.20b)$$

$$\mu_3 = (b - a)^3 \left[\frac{2\alpha\beta(\beta - \alpha)}{(\alpha + \beta + 2)(\alpha + \beta + 1)(\alpha + \beta)^3} \right] \quad (2.20c)$$

$$\mu_4 = (b-a)^4 \left[\frac{3\alpha\beta}{(\alpha+\beta+3)(\alpha+\beta+2)(\alpha+\beta+1)(\alpha+\beta)^4} \right] \quad (2.20d)$$

It is also possible to calculate the four beta parameters from the mean, variance, third and fourth central moments. The beta distribution shape parameters, α and β , may be recovered from the mean and first three central moments using the following relations [8]:

$$r = \frac{6(\beta_2 - \beta_1 - 1)}{6 + 3\beta_1 - 2\beta_2} \quad (2.21)$$

$$\alpha, \beta = \frac{1}{2}r \left\{ 1 \pm (r+2) \sqrt{\beta_1 \left\{ (r+2)^2 \beta_1 + 16(r+1) \right\}^{-1}} \right\} \quad (2.22)$$

where

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} \quad (2.23)$$

If $\mu_3 > 0$, then $\beta > \alpha$, otherwise $\alpha > \beta$.

Once the shape factors α and β are known, the endpoints a and b may be calculated by rearranging Equations (2.20a) and (2.20b) above to yield:

$$c = \sqrt{\mu_2 \frac{(\alpha+\beta+1)(\alpha+\beta)^2}{\alpha\beta}} \quad (2.24a)$$

$$a = \mu_1' - c \frac{\alpha}{(\alpha+\beta)} \quad (2.24b)$$

$$b = a + c \quad (2.24c)$$

This shows that a beta distribution is uniquely and completely characterized by its first four moments. The above result allows us to attempt to approximate a function of one or more independent beta random variables by simply calculating the first four moments of the result and fitting the corresponding beta distribution to this result. There is however no guarantee that a

beta distribution will serve as a good approximation to a function of independent beta distributions, since products, linear combinations, quotients, and powers of independent beta distributions are not in general beta distributions. It is experimentally tested in Chapter 5 whether products, sums, quotients, and powers of independent beta distributions can be well approximated by beta distributions.

3 Analytical Approximation Methods

This chapter addresses the problem of approximating linear combinations, products, quotients, and powers of independent beta variates. The methods described are intended to greatly simplify and quicken the process of performing mathematical operations with beta variates by avoiding having to perform convolution or Monte Carlo analysis (see Chapter 4). Each method enables calculation of the moments of a function of two or more random variables given only the moments of these random variables. Two separate methods are described for multiplication, division and exponentiation and one method for linear combinations of random variables. Section 3.1 describes methods developed by Springer in 1978 (see [16]), and Section 3.2 describes methods developed by McNichols in 1976 (see [11]).

Figure 3.1 shows the general method that is used to calculate the moments of two independent beta variables. The first step is to calculate the required moments or cumulants (see Section 3.1 for a treatment of cumulants) of the random variables X_1 and X_2 . Once these have been calculated, either one of Springer's or McNichols' methods is used to calculate the first four moments or cumulants of the resulting distribution. These can then be used to calculate the four beta parameters of the resulting approximation using the relations given in Section 2.2.

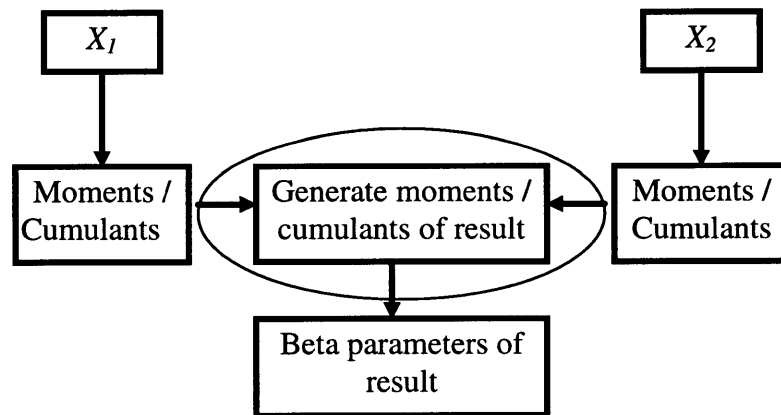


Figure 3.1: The general method used to approximate a function of two independent beta variables is represented here. The step inside the circle will differ depending on the functional form of the function to be approximated and the approximation method used (i.e. Springer's or McNichols').

3.1 Springer's Methods

In [16] Springer develops relations for exactly calculating the moments of a random variable that is a linear combination, product, or quotient of two other independent random variables as well as the moments of a random variable raised to a scalar power (exponentiation). As described in Section 3.1.1, the expression for the moments of a linear combination of independent random variables relies on a set of descriptive constants known as cumulants, which are derived from the logarithm of the characteristic function described in Section 2.1.3. Mellin transform theory is employed in Sections 3.1.2 and 3.1.3 to derive expressions for the moments of the product of two independent random variables and the moments of one random variable raised to a scalar power. Finally, expressions for the moments of the quotient and difference of two random variables are derived in Section 3.1.4 as special cases of multiplication and exponentiation, and linear combinations respectively.

3.1.1 Linear Combinations of Two Independent Random Variables

In order to calculate the moments of the random variable $Y = a + bX_1 + cX_2$ we will require a set of descriptive constants other than the moments, as the moments of Y cannot be readily calculated from the moments of X_1 and X_2 . This set of descriptive constants of a random variable are called cumulants and are derived using the characteristic function of a real random variable, which was described in Section 2.1.3. The characteristic function of the random variable $Y = X_1 + X_2$, $\phi(t)$, is simply the product of the characteristic functions, $\phi_1(t)$ and $\phi_2(t)$, of the two independent random variables, X_1 and X_2 , [13]:

$$\begin{aligned}\phi(t) &= E(e^{itY}) = E(e^{it(X_1+X_2)}) = E(e^{itX_1} \cdot e^{itX_2}) \\ &= E(e^{itX_1})E(e^{itX_2}) = \phi_1(t)\phi_2(t)\end{aligned}\tag{3.1}$$

By taking the natural logarithm of the characteristic function, an additive relationship can be established. It is therefore expedient to define the following function, known as the cumulant generating function:

$$\psi(t) = \ln(\phi(t))\tag{3.2}$$

The cumulant generating function of Y , $\psi(t)$, can now be expressed as the sum of the cumulant generating functions of X_1 and X_2 , $\psi_1(t)$ and $\psi_2(t)$:

$$\psi(t) = \psi_1(t) + \psi_2(t) \quad (3.3)$$

Similar to the way the moments were defined in terms of the characteristic function, we define the r^{th} cumulant of Y in terms of the r^{th} derivative of $\psi(t)$, the cumulant generating function of Y :

$$\kappa_r = i^{-r} \left. \frac{d^r \psi}{dt^r} \right|_{t=0} \quad (3.4)$$

Using the above relations, one can derive expressions for the cumulants of Y in terms of the moments of Y . This is done by Kendall and Stuart [9] for the cumulants up to order 10. For the purpose of this thesis only the first four cumulants of any given random variable will be needed. These are given in terms of the moments by:

$$\begin{aligned} \kappa_1 &= \mu'_1 \\ \kappa_2 &= \mu'_2 - \mu_1'^2 \\ \kappa_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3 \\ \kappa_4 &= \mu'_4 - 4\mu'_3\mu'_1 - 3\mu_2'^2 + 12\mu'_2\mu_1'^2 - 6\mu_1'^4 \end{aligned} \quad (3.5)$$

Manipulation of the above set of equations yields expressions for the moments in terms of the cumulants:

$$\begin{aligned} \mu'_1 &= \kappa_1 \\ \mu'_2 &= \kappa_2 + \kappa_1^2 \\ \mu'_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3 \\ \mu'_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4 \end{aligned} \quad (3.6)$$

The variance, third, and fourth central moments can also be expressed in terms of the cumulants:

$$\begin{aligned} \mu_2 &= \sigma^2 = \kappa_2 \\ \mu_3 &= \kappa_3 \\ \mu_4 &= \kappa_4 + 3\kappa_2^2 \end{aligned} \quad (3.7)$$

It is now possible to derive expressions for the first four cumulants of $Y = a + bX_1 + cX_2$ in terms of a, b, c , and the cumulants of X_1 and X_2 . The characteristic function of Y is given by [12]:

$$\phi(t) = E(e^{itY}) = E(\exp\{it(a + bX_1 + cX_2)\}) = e^{ita}(\phi_1(tb) \cdot \phi_2(tc)) \quad (3.8)$$

By taking the natural logarithm of $\phi(t)$ we obtain its cumulant generating function:

$$\psi(t) = ita + \phi_1(tb) + \phi_2(tc) \quad \begin{matrix} \psi(t) = ita + \ln(\phi_1(tb)) + \ln(\phi_2(tc)) \\ = ita + \ln(\phi_1(tb)) + \ln(\phi_2(tc)) \end{matrix} \quad (3.9)$$

Employing Equation (3.4), we find the following expressions for the first four cumulants of Y [11]:

$$\begin{bmatrix} \kappa_{1y} \\ \kappa_{2y} \\ \kappa_{3y} \\ \kappa_{4y} \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b \cdot \kappa_{1x_1} \\ b^2 \cdot \kappa_{2x_1} \\ b^3 \cdot \kappa_{3x_1} \\ b^4 \cdot \kappa_{4x_1} \end{bmatrix} + \begin{bmatrix} c \cdot \kappa_{1x_2} \\ c^2 \cdot \kappa_{2x_2} \\ c^3 \cdot \kappa_{3x_2} \\ c^4 \cdot \kappa_{4x_2} \end{bmatrix} \quad (3.10)$$

3.1.2 Multiplication of Two Independent Random Variables

It is necessary to use the Mellin Transform to derive an expression for the moments of $Y = X_1 X_2$ from the moments of X_1 and X_2 . If the integral:

$$\int_0^{\infty} x^{k-1} |f(x)| dx, \text{ where } x \geq 0 \quad (3.11)$$

converges for x and k real, then the Mellin transform of $f(x)$ exists and is given by [16]:

$$M_s(f(x)) = \int_0^{\infty} x^{s-1} f(x) dx \quad (3.12)$$

where s is complex. $f(x)$ can be recovered from $M_s(f(x))$ using the Mellin inversion integral:

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} M_s(f(x)) ds \quad (3.13)$$

where c is a real constant. The Mellin Transform of the density function of $Y = X_1X_2$ is given by:

$$M_s(g(y)) = E(Y^{s-1}) = E((X_1X_2)^{s-1}) \quad (3.14)$$

If X_1 and X_2 are independent, we may express the probability of the intersection of the events $X_1 \in A$ and $X_2 \in B$ as [13]:

$$\begin{aligned} P\{(X_1 \in A)(X_2 \in B)\} &= P\{(X_1, X_2) \in A \times B\} \\ &= \int \int_{A \times B} f(x_1, x_2) dx_1 dx_2 \\ &= \int_A f_1(x_1) dx_1 \int_B f_2(x_2) dx_2 \end{aligned} \quad (3.15)$$

Thus if X_1 and X_2 are independent, the Mellin Transform of their product is given by the product of their Mellin Transforms:

$$E((X_1X_2)^{s-1}) = E(X_1^{s-1})E(X_2^{s-1}) = M_s(f_1(x_1))M_s(f_2(x_2)) \quad (3.16)$$

From the definition of the Mellin transform given above and the definition of the r^{th} moment given in (2.3), it is clear that:

$$\mu'_r = E(X^r) = M_{r+1}(f(x)), \text{ where } x \geq 0 \quad (3.17)$$

Combining Equations (3.16) and (3.17), we find that the r^{th} moment of $Y = X_1X_2$ is given by:

$$\mu'_{ry} = \mu'_{rx_1} \mu'_{rx_2} \quad (3.18)$$

Fan [4] compared the first ten moments of products of several beta distributions obtained using the above approximation with the moments of products of the same beta distributions obtained analytically using Mellin transform methods and showed that the above approximation yields very good results.

3.1.3 Exponentiation of One Random Variable

The Mellin Transform is required to derive a general expression for the moments of the random variable $Y = X^m$. If Y has density function $g(y)$ and X has density function $f(x)$, then the Mellin transform of Y is given by [16]:

$$M_s(g(y)) = E(Y^{s-1}) = E\left((X^m)^{s-1}\right) = \int_0^{\infty} x^{m(s-1)} f(x) dx \quad (3.19)$$

From Equation (3.17), the r^{th} moment of Y (if it exists) is given by letting $s = r + 1$ in the above equation:

$$\mu'_{ry} = E(Y^r) = \int_0^{\infty} x^{mr} f(x) dx = \mu'_{(m \cdot r)x} \quad (3.20)$$

Thus, the r^{th} moment of $Y = X^m$ is given by the $m \cdot r^{\text{th}}$ moment of X .

If the density function of X is defined on the closed interval $[a, b]$, and $b > a > 0$, then the above integral will converge for all m and r , and we may write:

$$\mu'_{ry} = E(Y^r) = E(X^{mr}) = \mu'_{(m \cdot r)x} \quad (3.21)$$

3.1.4 Division and Subtraction of Two Random Variables

Both division and subtraction can be accomplished using the methods described in the previous three sections of this chapter. To find the moments of $Y = X_1 / X_2 = X_1(X_2)^{-1}$ we can use Equations (3.18) and (3.20):

$$\mu'_{ry} = \mu'_{rx_1} \mu'_{(-r)x_2} \quad (3.22)$$

Similarly, the cumulants of $Y = X_1 - X_2 = X_1 + (-1)X_2$ can be found from Equation (3.10):

$$\kappa_{ry} = \kappa_{rx_1} - \kappa_{rx_2} \quad (3.23)$$

3.2 McNichols' Methods

Using Springer's methods, exact expressions were found for the moments of a random variable Y which is a function of one or more other random variables. In [11] McNichols develops first and second order methods for approximating the moments of a random variable $Y = \varphi(X_1, X_2, \dots, X_n)$, where φ is any arbitrary function of real random variables. His first order method relies on finding a linear approximation to the function φ , while his second order method depends on finding a quadratic approximation thereof.

If a function $Y = \varphi(X_1, X_2, \dots, X_n)$ has continuous first partial derivatives for each X_i , then it can be approximated about the point $X^o = (X_1^o, X_2^o, \dots, X_n^o)$ by a first order polynomial Y_1 as follows:

$$Y_1 = \varphi(X_1, X_2, \dots, X_n) = \varphi(X^o) + \sum_{i=1}^n \left. \frac{\partial \varphi(X)}{\partial X_i} \right|_{X=X^o} (X_i - X_i^o) \quad (3.24)$$

Similarly, if a function $Y = \varphi(X_1, X_2, \dots, X_n)$ has continuous first and second partial derivatives for each X_i , then it can be approximated about the point $X^o = (X_1^o, X_2^o, \dots, X_n^o)$ by a second order polynomial Y_2 as follows:

$$\begin{aligned} Y_2 = \varphi(X_1, X_2, \dots, X_n) = & \varphi(X^o) + \sum_{i=1}^n \left. \frac{\partial \varphi(X)}{\partial X_i} \right|_{X=X^o} (X_i - X_i^o) \\ & + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \left. \frac{\partial^2 \varphi(X)}{\partial X_i \partial X_j} \right|_{X^o} \frac{(X_i - X_i^o)(X_j - X_j^o)}{2} \\ & + \sum_{i=1}^n \left. \frac{\partial^2 \varphi(X)}{\partial X_i^2} \right|_{X=X^o} \frac{(X_i - X_i^o)^2}{2} \end{aligned} \quad (3.25)$$

Since the density of a random variable will tend to be concentrated around its mean, the highest degree of accuracy will usually be obtained by applying the above approximations about the point $X^o = (\mu'_{1x_1}, \mu'_{1x_2}, \dots, \mu'_{1x_n})$ [13]. If Y is a function of one or more random variables whose density is not concentrated about the mean (such as a beta variate with α and $\beta < 1$, see Section 2.2), then we would not expect our approximation to be very good.

McNichols has applied (3.24) and (3.25) about the point $X^o = (\mu'_{1x_1}, \mu'_{1x_2}, \dots, \mu'_{1x_n})$ in order to find approximations to the cumulants of $Y = \varphi(X)$ as functions of the central moments of X [11]. His results are summarized in Table 3.1 and Table 3.2. For the case of linear combinations of independent random variables, McNichols' first and second order methods reduce to Springer's method (see Equation (3.10)). Expressions for the first four cumulants of the product and quotient of two independent random variables and those of a random variable raised to a scalar power have been derived for McNichols' first and second order methods and the results summarized in Table 3.3 through Table 3.8.

Cumulant of Y	Expression in terms of the central moments of X_1, X_2, \dots, X_n
κ_{1y}	$\varphi(\mu'_{1x_1}, \mu'_{1x_2}, \dots, \mu'_{1x_n})$
κ_{2y}	$\sum_{i=1}^n \left(\frac{\partial \varphi(X)}{\partial X_i} \Big _{\mu'_{1x}} \right)^2 \mu_{2x_i}$
κ_{3y}	$\sum_{i=1}^n \left(\frac{\partial \varphi(X)}{\partial X_i} \Big _{\mu'_{1x}} \right)^3 \mu_{3x_i}$
κ_{4y}	$\sum_{i=1}^n \left(\frac{\partial \varphi(X)}{\partial X_i} \Big _{\mu'_{1x}} \right)^4 (\mu_{4x_i} - 3\mu_{2x_i}^2)$

Table 3.1: The cumulants of $Y = \varphi(X_1, X_2, \dots, X_n)$ as a function of the central moments of X_1, X_2, \dots, X_n for McNichols' first order approximation (see Equation (3.22))

κ_{1y}	$\varphi(\mu'_{1x_1}, \mu'_{1x_2}, \dots, \mu'_{1x_n}) + \frac{1}{2} \sum_{i=1}^n \left(\frac{\partial^2 \varphi(X)}{\partial X_i^2} \Big _{\mu'_{ix}} \right) \mu_{2x_i}$
κ_{2y}	$\sum_{i=1}^n \left(\frac{\partial \varphi(X)}{\partial X_i} \Big _{\mu'_{ix}} \right)^2 \mu_{2x_i} + \frac{1}{4} \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \left(\frac{\partial^2 \varphi(X)}{\partial X_i \partial X_j} \Big _{\mu'_{ix}} \right)^2 \mu_{2x_i} \mu_{2x_j}$ $+ \frac{1}{4} \sum_{i=1}^n \left(\frac{\partial \varphi^2(X)}{\partial X_i^2} \Big _{\mu'_{ix}} \right)^2 (\mu_{4x_i} - 3\mu_{2x_i}^2)$
κ_{3y}	$\sum_{i=1}^n \left(\frac{\partial \varphi(X)}{\partial X_i} \Big _{\mu'_{ix}} \right)^3 \mu_{3x_i} + \frac{1}{8} \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \left(\frac{\partial^2 \varphi(X)}{\partial X_i \partial X_j} \Big _{\mu'_{ix}} \right)^3 \mu_{3x_i} \mu_{3x_j}$ $+ \frac{1}{8} \sum_{i=1}^n \left(\frac{\partial \varphi^2(X)}{\partial X_i^2} \Big _{\mu'_{ix}} \right)^3 (\mu_{6x_i} - 3\mu_{4x_i} \mu_{2x_i} + 4\mu_{2x_i}^2)$
κ_{4y}	$\sum_{i=1}^n \left(\frac{\partial \varphi(X)}{\partial X_i} \Big _{\mu'_{ix}} \right)^4 (\mu_{4x_i} - 3\mu_{2x_i}^2)$ $+ \frac{1}{16} \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \left(\frac{\partial^2 \varphi(X)}{\partial X_i \partial X_j} \Big _{\mu'_{ix}} \right)^4 (\mu_{4x_i} \mu_{4x_j} - 3\mu_{2x_i}^2 \mu_{2x_j}^2)$ $+ \frac{1}{16} \sum_{i=1}^n \left(\frac{\partial \varphi^2(X)}{\partial X_i^2} \Big _{\mu'_{ix}} \right)^4 (\mu_{8x_i} - 4\mu_{6x_i} \mu_{2x_i} + 12\mu_{4x_i} \mu_{2x_i}^2 - 6\mu_{2x_i}^4 - 3\mu_{4x_i}^2)$

Table 3.2: The cumulants of $Y = \varphi(X_1, X_2, \dots, X_n)$ as a function of the central moments of X_1, X_2, \dots, X_n for McNichols' second order approximation (see Equation (3.23))

Cumulant of $Y = X_1X_2$

Expression in terms of the central moments of X_1 and X_2

κ_{1y}	$\mu'_{1x_1}\mu'_{1x_2}$
κ_{2y}	$\mu'^2_{1x_1}\mu_{2x_2} + \mu'^2_{1x_2}\mu_{2x_1}$
κ_{3y}	$\mu'^3_{1x_1}\mu_{3x_2} + \mu'^3_{1x_2}\mu_{3x_1}$
κ_{4y}	$\mu'^4_{1x_1}(\mu_{4x_2} - 3\mu_{2x_2}^2) + \mu'^4_{1x_2}(\mu_{4x_1} - 3\mu_{2x_1}^2)$

Table 3.3: The cumulants of $Y = X_1X_2$ as a function of the central moments of X_1 and X_2 for McNichols' first order approximation

Cumulant of $Y = X_1X_2$

Expression in terms of the central moments of X_1 and X_2

κ_{1y}	$\mu'_{1x_1}\mu'_{1x_2}$
κ_{2y}	$\mu'^2_{1x_1}\mu_{2x_2} + \mu'^2_{1x_2}\mu_{2x_1} + \frac{1}{2}\mu_{2x_1}\mu_{2x_2}$
κ_{3y}	$\mu'^3_{1x_1}\mu_{3x_2} + \mu'^3_{1x_2}\mu_{3x_1} + \frac{1}{4}\mu_{3x_1}\mu_{3x_2}$
κ_{4y}	$\mu'^4_{1x_1}(\mu_{4x_2} - 3\mu_{2x_2}^2) + \mu'^4_{1x_2}(\mu_{4x_1} - 3\mu_{2x_1}^2)$ $+ \frac{1}{8}(\mu_{4x_1}\mu_{4x_2} - 3\mu_{2x_1}^2\mu_{2x_2}^2)$

Table 3.4: The cumulants of $Y = X_1X_2$ as a function of the central moments of X_1 and X_2 for McNichols' second order approximation

Cumulant of $Y = X_1/X_2$ Expression in terms of the central moments of X_1 and X_2

κ_{1y}	$\frac{\mu'_{1x_1}}{\mu'_{1x_2}}$
κ_{2y}	$\frac{\mu_{2x_1}}{\mu'^2_{1x_2}} + \left(\frac{\mu'_{1x_1}}{\mu'^2_{1x_2}}\right)^2 \mu_{2x_2}$
κ_{3y}	$\frac{\mu_{3x_1}}{\mu'^3_{1x_2}} - \left(\frac{\mu'_{1x_1}}{\mu'^2_{1x_2}}\right)^3 \mu_{3x_2}$
κ_{4y}	$\frac{\mu_{4x_1} - 3\mu^2_{2x_1}}{\mu'^4_{1x_2}} + \left(\frac{\mu'_{1x_1}}{\mu'^2_{1x_2}}\right)^4 (\mu_{4x_2} - 3\mu^2_{2x_2})$

Table 3.5: The cumulants of $Y = X_1/X_2$ as a function of the central moments of X_1 and X_2 for McNichols' first order approximationCumulant of $Y = X_1/X_2$ Expression in terms of the central moments of X_1 and X_2

κ_{1y}	$\frac{\mu'_{1x_1}}{\mu'_{1x_2}} + \left(\frac{\mu'_{1x_1}}{\mu'^3_{2x_2}}\right) \mu_{2x_2}$
κ_{2y}	$\frac{\mu_{2x_1}}{\mu'^2_{1x_2}} + \left(\frac{\mu'_{1x_1}}{\mu'^2_{1x_2}}\right)^2 \mu_{2x_2} + \frac{1}{2} \frac{\mu_{2x_1}}{\mu'^4_{1x_2}} \mu_{2x_2} + \left(\frac{\mu'_{1x_1}}{\mu'^3_{1x_2}}\right)^2 (\mu_{4x_2} - 3\mu^2_{2x_2})$
κ_{3y}	$\frac{\mu_{2x_1}}{\mu'^3_{1x_2}} - \left(\frac{\mu'_{1x_1}}{\mu'^2_{1x_2}}\right)^3 \mu_{3x_2} - \frac{1}{4} \frac{\mu_{3x_1}}{\mu'^6_{1x_2}} \mu_{3x_2} + \left(\frac{\mu'_{1x_1}}{\mu'^3_{1x_2}}\right)^3 (\mu_{6x_2} - 3\mu_{4x_2} \mu_{2x_2} + 4\mu^3_{2x_2})$
κ_{4y}	$\frac{\mu_{4x_1} - 3\mu^2_{2x_1}}{\mu'^4_{1x_2}} + \left(\frac{\mu'_{1x_1}}{\mu'^2_{1x_2}}\right)^4 (\mu_{4x_2} - 3\mu^2_{2x_2}) + \frac{1}{8} \frac{\mu_{4x_1} \mu_{4x_2} - 3\mu^2_{2x_1} \mu^2_{2x_2}}{\mu'^8_{1x_2}}$ $+ \left(\frac{\mu'_{1x_1}}{\mu'^3_{1x_2}}\right)^4 (\mu_{8x_2} - 4\mu_{6x_2} \mu_{2x_2} + 12\mu_{4x_2} \mu^2_{2x_2} - 6\mu^4_{2x_2} - 3\mu^2_{4x_2})$

Table 3.6: The cumulants of $Y = X_1/X_2$ as a function of the central moments of X_1 and X_2 for McNichols' second order approximation

Cumulant of $Y = X^m$	Expression in terms of the central moments of X
-----------------------	---

K_{1y}	μ'_{1x_1}
K_{2y}	$m^2 \mu'_{1x}{}^{2(m-1)} \mu_{2x}$
K_{3y}	$m^3 \mu'_{1x}{}^{3(m-1)} \mu_{3x}$
K_{4y}	$m^4 \mu'_{1x}{}^{4(m-1)} (\mu_{4x} - 3\mu_{2x}^2)$

Table 3.7: The cumulants of $Y = X^m$ as a function of the central moments of X_I for McNichols' first order approximation

Cumulant of $Y = X^m$	Expression in terms of the central moments of X
K_{1y}	$\mu'_{1x}{}^m + \frac{1}{2} m(m-1) \mu'_{1x}{}^{m-2} \mu_{2x}$
K_{2y}	$m^2 \mu'_{1x}{}^{2(m-1)} \mu_{2x} + \frac{1}{4} (m(m-1) \mu'_{1x}{}^{(m-2)})^2 (\mu_{4x} - 3\mu_{2x}^2)$
K_{3y}	$m^3 \mu'_{1x}{}^{3(m-1)} \mu_{3x} + \frac{1}{8} (m(m-1) \mu'_{1x}{}^{(m-2)})^3 (\mu_{6x} - 3\mu_{4x} \mu_{2x} + 4\mu_{2x}^3)$
K_{4y}	$m^4 \mu'_{1x}{}^{4(m-1)} \mu_{4x} + \frac{1}{16} (m(m-1) \mu'_{1x}{}^{(m-2)})^4 (\mu_{8x} - 4\mu_{6x} \mu_{2x} + 12\mu_{4x} \mu_{2x}^2 - 6\mu_{2x}^4 - 3\mu_{4x}^2)$

Table 3.8: The cumulants of $Y = X^m$ as a function of the central moments of X_I for McNichols' second order approximation

4 Numerical Methods

This chapter explores methods for finding numerical approximations to the density functions of the sums, differences, products, quotients, and powers of random variables. Given $Y = \varphi(X_1, X_2)$, or $Y = \varphi(X)$ the result of all of the methods described is to generate the vectors \mathbf{g} and \mathbf{y} , which represent $g(y)$, the density function of Y , and y sampled at n points. Section 4.1 introduces the Fourier and Mellin convolution integrals associated with addition, subtraction, multiplication, and division of two independent random variables as well an analytic expression for the density function of a random variable raised to a scalar power. Algorithms to find numerical approximations to these analytical expressions are developed in Section 4.2. Finally, Monte Carlo methods are developed in Section 4.3 to determine approximations to \mathbf{g} and \mathbf{y} for any arbitrary function φ .

4.1 Analytical Solutions to Functions of Random Variables

Although this thesis intentionally avoids analytic solution of expressions associated with functions of independent random variables, these expressions form the basis of the numerical methods for finding numerical approximations to sums, differences, products, quotients and powers of random variables described in Section 4.2. The convolution integrals associated with sums, differences, products, and quotients of two random variables are therefore derived in Section 4.1.1, and an expression for the density function of one random variable raised to a scalar power is introduced in Section 4.1.2.

4.1.1 Analytical Convolution of Two Independent Random Variables

Springer [16] gives analytic expressions in the form of convolution integrals for products, sums, quotients, and differences of two independent random variables.

In the case of $Y = X_1 + X_2$ the density function of Y , $g(y)$, as a function of the density functions of X_1 and X_2 , $f_1(x_1)$ and $f_2(x_2)$ is given by the Fourier convolution integral:

$$g(y) = \int_{-\infty}^{\infty} f_1(y - x_2) \cdot f_2(x_2) dx_2 \quad (4.1)$$

In order to derive this result let:

$$\begin{aligned} x_1 &= y - x_2 \\ x_2 &= x_2 \end{aligned} \quad (4.2)$$

Since X_1 and X_2 are independent, their joint probability element is given by:

$$f(x_1, x_2) dx_1 dx_2 = f(x_1) f(x_2) dx_1 dx_2 \quad (4.3)$$

Using (4.2), this becomes:

$$g(y, x_2) dx_2 dy = f_1(y - x_2) f_2(x_2) J dx_2 dy \quad (4.4)$$

Where J is the Jacobian of the transformation given in (4.2):

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial x_2} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial x_2} \end{vmatrix} \\ &= \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1 \end{aligned} \quad (4.5)$$

The variable x_2 can be integrated out of (4.4) to yield the density function of $Y = X_1 + X_2$, given in (4.1).

Employing the above method and the following transformation:

$$\begin{aligned} x_1 &= y + x_2 \\ x_2 &= x_2 \end{aligned} \quad (4.6)$$

the convolution integral $Y = X_1 - X_2$ is found to be:

$$g(y) = \int_{-\infty}^{\infty} f_1(y + x_2) \cdot f_2(x_2) dx_2 \quad (4.7)$$

In the case of $Y = X_1 X_2$, the transformation:

$$\begin{aligned} x_1 &= \frac{y}{x_2} \\ x_2 &= x_2 \end{aligned} \quad (4.8)$$

yields the Mellin integral for the product of two random variables:

$$g(y) = \int_{-\infty}^{\infty} \frac{1}{x_2} f_1\left(\frac{y}{x_2}\right) \cdot f_2(x_2) dx_2 \quad (4.9)$$

Finally, the transformation:

$$\begin{aligned} x_1 &= yx_2 \\ x_2 &= x_2 \end{aligned} \quad (4.10)$$

yields the Mellin integral for the random variable $Y = \frac{X_1}{X_2}$:

$$g(y) = \int_{-\infty}^{\infty} x_2 f_1(yx_2) \cdot f_2(x_2) dx_2 \quad (4.11)$$

The integrals given in Equations (4.1) and (4.7) may be solved using Laplace transform methods and the integrals given in Equations (4.9) and (4.11) using Mellin transform methods. Unfortunately, these analytic solutions can be very difficult to compute and do not lend themselves well to implementation on a computer. Dennis [2] finds an analytic expression for the product of three beta variables using Mellin transform methods, which contains 23 terms and still only serves as an approximation to the actual solution.

4.1.2 Analytical Exponentiation of One Random Variable

Numerical evaluation of a convolution integral is not necessary in order to find the density of a function of one random variable. Given a real random variable X and an arbitrary function φ , the density function of $Y = \varphi(X)$ is [9]:

$$g(y) = f(\varphi^{-1}(y)) \frac{d\varphi^{-1}}{dy} \quad (4.12)$$

For the special case of $Y = X^m$:

$$\varphi^{-1}(y) = y^{1/m} \quad (4.13)$$

Substituting (4.13) into (4.12), we find an expression for the density function of $Y = X^m$:

$$g(y) = \frac{1}{|m|} f(y^{1/m}) y^{1/m-1}, \quad y \geq 0 \quad (4.14)$$

4.2 Direct Numerical Approximation to Functions of Random Variables

Since an exact analytic solution to the equations introduced in the previous section may be tedious to compute and is usually not required, it is advantageous to solve them numerically. Iqbal [18] has found a methodology to numerically invert Mellin integrals based on the expansion of Laguerre polynomials, but his method has been left as a topic for further research. Section 4.2.1 develops methods for obtaining numeric approximations to the convolution integrals in Section 4.1.1 based on sampling the density functions of $f_1(x_1)$ and $f_2(x_2)$. Section 4.2.2 introduces a simple algorithm for directly calculating a numerical approximation to Equation (4.14).

4.2.1 Numerical Convolution of Two Independent Random Variables

This section describes the numerical techniques used to compute the vectors \mathbf{g} and \mathbf{y} , corresponding respectively to sampled versions of $g(y)$ and y given in Equations (4.1) and (4.7).

If the two density functions $f_1(x_1)$ and $f_2(x_2)$ are defined on the closed intervals $[a_1, b_1]$ and $[a_2, b_2]$, y will be defined on the closed interval $[c, d]$, where:

$$\begin{aligned} c &= a_1 + a_2 \\ d &= b_1 + b_2 \end{aligned} \tag{4.15}$$

for the case of addition of X_1 and X_2 ;

$$\begin{aligned} c &= a_1 - b_2 \\ d &= b_1 - a_2 \end{aligned} \tag{4.16}$$

for the case of subtraction of X_1 and X_2 ;

$$\begin{aligned} c &= a_1 \cdot a_2 \\ d &= b_1 \cdot b_2 \end{aligned} \tag{4.17}$$

for the case of multiplication of X_1 and X_2 ; and:

$$\begin{aligned} c &= a_1 / b_2 \\ d &= b_1 / a_2 \end{aligned} \tag{4.18}$$

for the case of division of X_1 and X_2 .

The n -element vector y may be obtained by sampling y at n evenly spaced points between c and d .

In order to find g for the case of the sum of two random variables, each of the density functions, $f_1(x_1)$ and $f_2(x_2)$, associated with the random variables X_1 and X_2 is first sampled at $\frac{n}{2}$ evenly spaced points, yielding the $\frac{n}{2}$ -element vectors f_1 and f_2 . If we let $n = 2^i$ for some positive integer i , an n point radix-2 Fast Fourier Transform (see [10]) may be performed on each of the vectors, yielding the n -element vectors \hat{f}_1 and \hat{f}_2 . We then generate the vector \hat{g} , whose i^{th} element is given by:

$$\hat{g}_i = \hat{f}_{1i} \cdot \hat{f}_{2i}, \quad i = 1, \dots, n \tag{4.19}$$

The n -element vector \mathbf{g} , representing $g(x)$ sampled at n evenly spaced points, is obtained by taking the n point radix-2 Inverse Fast Fourier Transform of $\hat{\mathbf{g}}$.

The procedure for finding \mathbf{g} for the difference of X_1 and X_2 is the same as that for addition except that the vector \mathbf{f}_2 must be reversed before its Fast Fourier Transform is taken such that (assuming vector indices $i = 1, \dots, n$):

$$\begin{aligned} f'_{2i} &= f_{2(n+1-i)}, \quad i = 1, \dots, n \\ \hat{\mathbf{f}}_2 &= \text{FFT}\{\mathbf{f}'_2\} \end{aligned} \quad (4.20)$$

Since Fourier Transform methods cannot be used to find numeric approximations to the convolution integrals given in Equations (4.9) and (4.11) for $Y = X_1 X_2$ and $Y = X_1 / X_2$, respectively, a novel algorithm was developed by this author to find their approximations. Given the n -element vectors \mathbf{y} and \mathbf{x}_2 , representing sampled versions of y and x_2 , the procedure given immediately below is used to obtain each point \mathbf{g}_i of the n -element vector \mathbf{g} , corresponding to $g(y)$ for the case of multiplication of two random variables.

1. The lower and upper limits of integration, a and b , of the Mellin integral may be found by:

$$\begin{aligned} a &= \max\left\{\frac{y_l}{a_1}, a_2\right\} \\ b &= \min\left\{\frac{y_l}{b_1}, b_2\right\} \end{aligned} \quad (4.21)$$

2. The lower and upper indices of the vector \mathbf{x}_2 , j_l and j_u , may be found using the relations (assuming vector indices from $j = 1, \dots, n$):

$$\begin{aligned} j_l &= \text{round}\left\{\frac{a - a_2}{b_2 - a_2}(n - 1) + 1\right\} \\ j_u &= \text{round}\left\{\frac{b - a_2}{b_2 - a_2}(n - 1) + 1\right\} \end{aligned} \quad (4.22)$$

3. Finally, each \mathbf{g}_i may be found from the following relation:

$$\mathbf{g}_i = \frac{1}{n} \sum_{j=j_l}^{j_u} \frac{1}{x_{2j}} f_1\left(\frac{y_i}{x_{2j}}\right) f_2(x_{2j}), \quad i = 1, \dots, n \quad (4.23)$$

In order to find g for the case of division of two random variables, the procedure given below is used.

1. The lower and upper limits of integration, a and b , of the Mellin integral may be found by:

$$\begin{aligned} a &= \max\left\{\frac{a_1}{y_i}, a_2\right\} \\ b &= \min\left\{\frac{b_1}{y_i}, b_2\right\} \end{aligned} \quad (4.24)$$

2. The lower and upper indices of the vector x_2 , j_l and j_u , may be found using (4.22).

3. Finally, each g_i may be found from the following relation:

$$g_i = \frac{1}{n} \sum_{j=j_l}^{j_u} x_{2j} f_1(y_i x_{2j}) f_2(x_{2j}), \quad i = 1, \dots, n \quad (4.25)$$

4.2.2 Numerical Exponentiation of One Random Variable

Given Equation (4.14), we may proceed to find a numeric approximation for $g(y)$. If x is defined on the closed interval $[a, b]$, then y will be defined on the interval $[a^m, b^m]$ if $m \geq 0$ and $[b^m, a^m]$ if $m < 0$. By sampling y at n evenly spaced points, we obtain the n -element vector y . Employing the previous Equation (4.14), g_i can be found:

$$g_i = \frac{1}{|m|} f(y_i^{1/m}) y_i^{1/m-1}, \quad i = 1, \dots, n \quad (4.26)$$

4.3 Monte Carlo Approximation of Functions of Random Variables

Monte Carlo analysis offers an alternative to convolution for calculating $Y = \varphi(X_1, X_2)$, where φ is any arbitrary real function [7]. If X_1 and X_2 are defined over two closed intervals, let y be the n -element vector of evenly spaced points between a and b , where a and b are given by:

$$\begin{aligned} a &= \min(\varphi(x_1, x_2)) \\ b &= \max(\varphi(x_1, x_2)) \end{aligned} \quad (4.27)$$

A pseudo-random number generator is then used to generate two p -element vectors u_1 and u_2 , where $p > n$, whose elements are uniformly distributed between 0 and 1. Each element of these two vectors is then mapped to a point in its respective distribution, resulting in two new vectors \mathbf{x}'_1 and \mathbf{x}'_2 . A description of this mapping for the case of generalized beta variates can be found in [1]. The i^{th} element of the p -element vector \mathbf{y}' is calculated from \mathbf{x}'_1 and \mathbf{x}'_2 by:

$$y'_i = \varphi(x'_{1i}, x'_{2i}), \quad i = 1, \dots, p \quad (4.28)$$

We may now subdivide the interval $[a, b]$ into n equally-sized disjoint intervals and form the n -element vector \mathbf{h} by counting the number of points of \mathbf{y}' that fall into each of these intervals. The i^{th} element of \mathbf{g} , the vector approximation to the density function of Y , is obtained through:

$$g_i = \frac{nh_i}{p} \quad (4.29)$$

The Monte Carlo method described above has the advantage over the direct numerical methods described in the previous two sections that it generalizes easily to any functional form, φ .

5 Implementation and Experimental Results

The previous three chapters have laid the theoretical groundwork for performing arithmetic operations on beta random variables. In this chapter, practical implementation of the numerical and analytical algorithms developed, as well as the concrete results they produce is discussed. Section 5.1 addresses the problem of generating sharp beta density functions numerically from their beta parameters on digital computers with finite floating point limits. The chi-squared test, which is used to evaluate the accuracy of all of the approximation methods under study, and the least squares test, which is used only to evaluate the accuracy of Springer's and McNichols' approximations, are the topic of Section 5.2. The various numerical methods introduced in Chapter 4 are compared in Section 5.3 with respect to their relative speeds and the smoothness of the density functions they produce. Implementation of Springer's and McNichols' analytical approximation methods, which were discussed in Chapter 3, is addressed along with a comparison of the different methods in terms of the accuracies of the approximations they produce in Section 5.4. Finally, Section 5.5 compares the execution times of the analytical and numerical routines under investigation. All of the routines described were implemented using MATLAB™ [10] on a 200MHz PC with a Pentium™ processor.

5.1 Calculation of a Beta Density Function from its Beta Parameters

In order to implement the numerical methods of Chapter 4 and test the analytic methods of Chapter 3 on beta density functions, an efficient and accurate means of generating sampled versions of these density functions is required. This section describes two methods for calculating the vector f , representing a sampled version of the generalized beta density function $f(x)$ described in Section 2.2. The first method is the more simple of the two and works correctly as long as the beta shape parameters of f and the interval upon which f is defined are small enough. If we wish to define f over a large interval or require a sharp, unimodal density function, this first method may not work due to the floating point limits of a digital computer. The second method described circumvents the floating point limits of digital computers by avoiding excessively small or large intermediate results and taking advantage of the fact that the integral of any probability density function over its entire domain must equal 1 (see Equation (2.2c)).

If $f(x)$ is defined on the closed interval $[a,b]$, we may define the vector \mathbf{x} to be the set of n evenly spaced points between a and b . When the beta shape parameters of f are sufficiently small and the interval upon which f is defined is sufficiently narrow, the vector \mathbf{f} , representing f sampled at n evenly spaced points, may be calculated directly by:

$$f_i = \frac{1}{\underbrace{(b-a)^{\alpha+\beta-1} \Gamma(\alpha)\Gamma(\beta)}_k} (x_i - a)^{\alpha-1} (b - x_i)^{\beta-1} \quad i = 1, \dots, n \quad (5.1)$$

Due to the fact that any probability density function must integrate to 1, k in the above equation is given by:

$$k = \left[\int_a^b (x-a)^{\alpha-1} (b-x)^{\beta-1} dx \right]^{-1} \quad (5.2)$$

When α , β , or $(b-a)$ are sufficiently large, k can exceed the upper floating point limit ($\approx 1.78 \times 10^{308}$ for purposes of this thesis) or the lower floating point limit ($\approx 2.2 \times 10^{-308}$) of the computer being employed. This problem can be circumvented by recognizing that:

$$f_i = k \left[(x_i - a)^{\frac{\alpha-1}{d}} (b - x_i)^{\frac{\beta-1}{d}} \right]^d, \quad i = 1, \dots, n \quad (5.3)$$

k can be neglected since it represents only a scale factor, and since it is known that any density function must integrate to 1 over its entire domain. If two scalar quantities, c and d , can be found such that when:

$$h_i = \left[c(x_i - a)^{\frac{\alpha-1}{d}} (b - x_i)^{\frac{\beta-1}{d}} \right]^d, \quad i = 1, \dots, n \quad (5.4)$$

is trapezoidally integrated with respect to \mathbf{x} over the interval $[a,b]$, the resulting scalar s exceeds neither floating point limit, the desired result \mathbf{f} can be calculated by:

$$f_i = \frac{h_i}{s}, \quad i = 1, \dots, n \quad (5.5)$$

c is determined by letting:

$$p = \frac{\alpha - 1}{d}, q = \frac{\beta - 1}{d} \quad (5.6)$$

$$c = \frac{1}{(b - a)^{p+q+1}} \frac{\Gamma(p+q+2)}{\Gamma(p+1)\Gamma(q+1)}$$

and finding the smallest positive integer d such that c is within the floating point limits. Since we know that when:

$$h'_i = \left[c(x_i - a)^{\frac{\alpha-1}{d}} (b - x_i)^{\frac{\beta-1}{d}} \right]^d, \quad i = 1, \dots, n \quad (5.7)$$

is trapezoidally integrated with respect to x , the result is 1, h will not exceed the floating point limits for sufficiently small values of d .

Letting $a = 100$, $b = 200$, $n = 1000$, and $\gamma = \alpha = \beta$, it was found that the maximum value of γ for which f could be calculated directly using Equation (5.1) was about 75, whereas the algorithm described above yielded valid beta density functions for values of γ in excess of 20,000. Figure 5.1 shows plots of $f_b = (x; 2 \times 10^4, 2 \times 10^4, 100, 200)$ and $f_b = (x; 75, 75, 100, 200)$ for comparison. These plots show clearly how the above algorithm allows for calculation of beta density functions with much sharper peaks than could be calculated using the direct method alone.

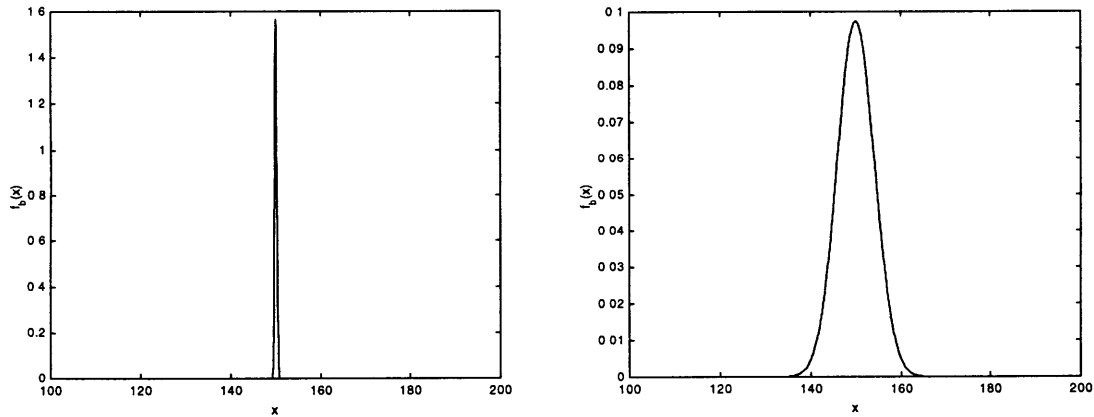


Figure 5.1: Plots of $f_b = (x; 2 \times 10^4, 2 \times 10^4, 100, 200)$ (left) and $f_b = (x; 75, 75, 100, 200)$ (right) are shown above. Note how much sharper the density function on the left is compared to the density function on the right.

5.2 Chi-Squared and Least-Squares Tests

All of the analytical approximation and numerical methods described in Chapters 3 and 4 were tested using a chi-squared (χ^2) test in order to determine their respective accuracies. In addition to the chi-squared test, the analytical approximation methods of Chapter 3 were also subjected to a least-squares test.

A flowgraph of the routine used to perform the chi-squared test is shown in Figure 5.2. $n = 1099$ points are sampled from each input beta distribution X_i . Each of these n points from the beta distribution $f_b(x; \alpha, \beta, a, b)$ was generated using Cheng's algorithm BA, described in [1], in the following manner:

1. Let $c = \alpha + \beta$. If $\min(\alpha, \beta) \leq 1$, let $d = \max(\alpha^{-1}, \beta^{-1})$; otherwise let $d = \sqrt{(c-2)/(2\alpha\beta-c)}$. Let $v = \alpha + d^{-1}$.
2. Generate the random numbers U_1 and U_2 such that $0 \leq U_1, U_2 \leq 1$ and let $V = d \log \sqrt{U_1 / (1 - U_1)}$. Let $W = \alpha e^V$.
3. If $c \log\{c / (\beta + W)\} + vV - \log(4) < \log(U_1^2 U_2)$, return to 2.
4. Otherwise deliver $a + \frac{(b-a)W}{(\beta+W)}$.

Given the functional form of $Y = \varphi(X_1, X_2)$ or $Y = \varphi(X)$, n points from Y can be generated. An approximating distribution to Y is then produced using one of the numerical or analytical approximation methods described in Chapters 3 and 4. A chi-squared test with a significance level of $\gamma = 0.01$ is then performed to determine the likelihood that the set of n points from the true distribution, Y , could have come from the distribution obtained through numerical or analytical approximation. $\gamma = 0.01$ was chosen as opposed to $\gamma = 0.05$ or $\gamma = 0.10$ because it was found that almost all distributions that passed a chi-squared test at this significance level provided excellent approximations to Y . The chi-squared test is accomplished by dividing the approximation to Y into k equiprobable intervals. Greenwood and Nikulin suggest selecting k such that [6]:

$$k \leq \min\left(\frac{1}{\gamma}, \log n\right) = 7 \quad (5.8)$$

The n randomly generated points from the true distribution of Y are grouped into these k equiprobable intervals. The result is the k -element vector \mathbf{p} . The chi-squared metric is then given by:

$$\chi^2 = \frac{k}{n} \sum_{i=1}^k (p_i)^2 - n \quad (5.9)$$

A table of the chi-squared distribution may be used to determine at what level of significance the approximation to Y matches the true distribution of Y , given χ^2 and $k - 1$ degrees of freedom. If this level of significance is greater than $\gamma = 0.01$, the approximation has passed the test, otherwise it has failed.

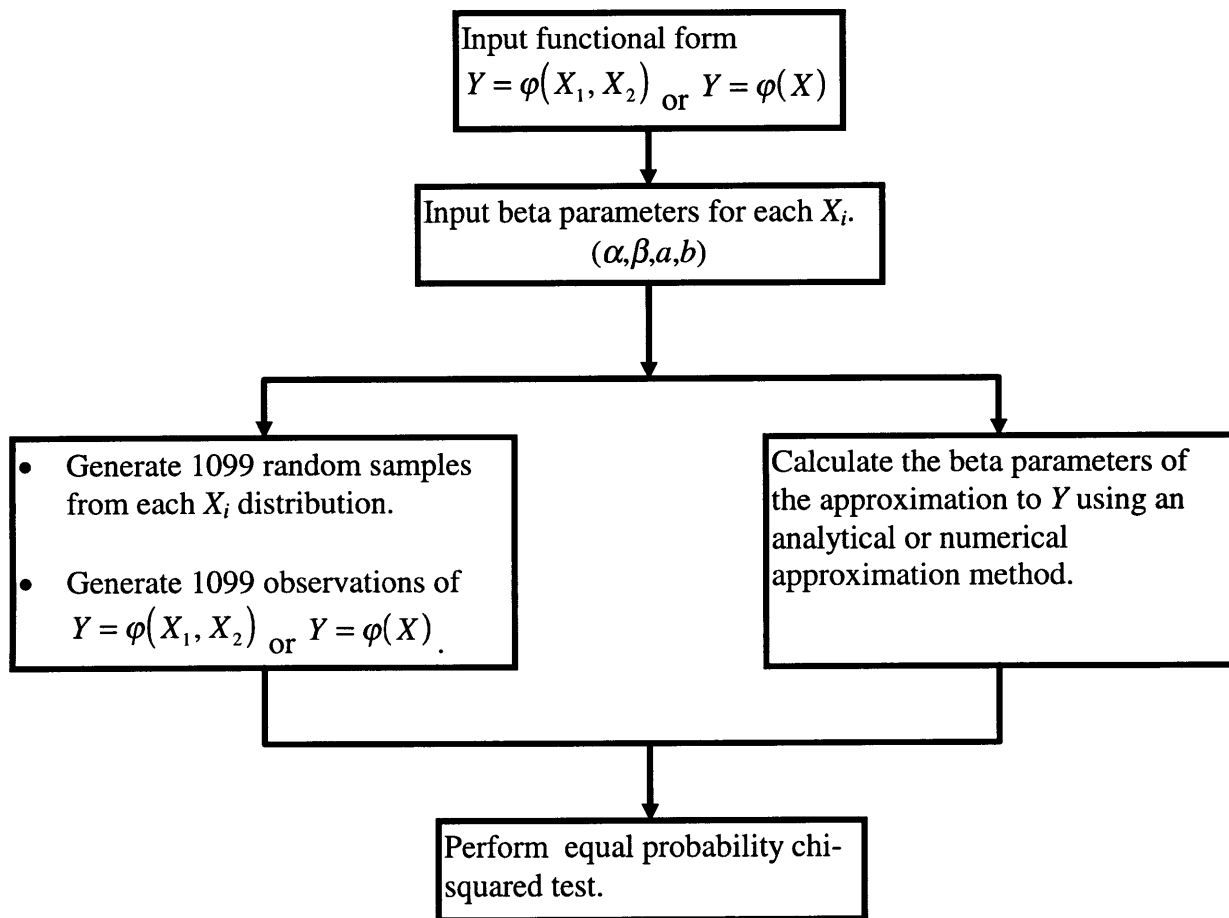


Figure 5.2: A flowgraph of the chi-squared routine used to test each all approximation routines is shown above.

A flowgraph of the routine used to perform the least-squares test is shown in Figure 5.3. Let g' be a sampled version of the 'true' density function of $Y = \varphi(X_1, X_2)$ or $Y = \varphi(X)$, calculated using one of the numerical methods described in Section 4.2. Generate the p -element vector y by sampling y at p evenly spaced points. The beta parameters, (α, β, a, b) , of the approximation to Y are calculated using one of either Springer's or McNichols' methods. If a' is the minimum of the vector y and b' its maximum value, let:

$$\begin{aligned} c &= \max(a, a') \\ d &= \min(b, b') \end{aligned} \tag{5.10}$$

If necessary, g' and y are truncated so that they are defined only on the interval $[c, d]$. Using this n -element truncated version of y , the beta shape parameters, α and β , and the algorithm described in Section 5.1, g , the n -element sampled version of the approximation to Y , is calculated. The result of the least-squares test is then given by [20]:

$$s = \frac{1}{n-1} \sum_{i=1}^n (g'_i - g_i)^2 \tag{5.11}$$

5.3 Implementation and Comparison of Numerical Methods

This section examines implementation of the numerical methods described in the previous chapter and compares these routines on the basis of their respective execution times and the relative smoothness of the curves generated by them. The convolution methods introduced in Section 4.2.1. are discussed in Section 5.3.1. The simple case of numerical exponentiation of one random variable to a scalar power is briefly discussed in Section 5.3.2. Section 5.3.3 explores implementation of the Monte Carlo methods developed in Section 4.3.

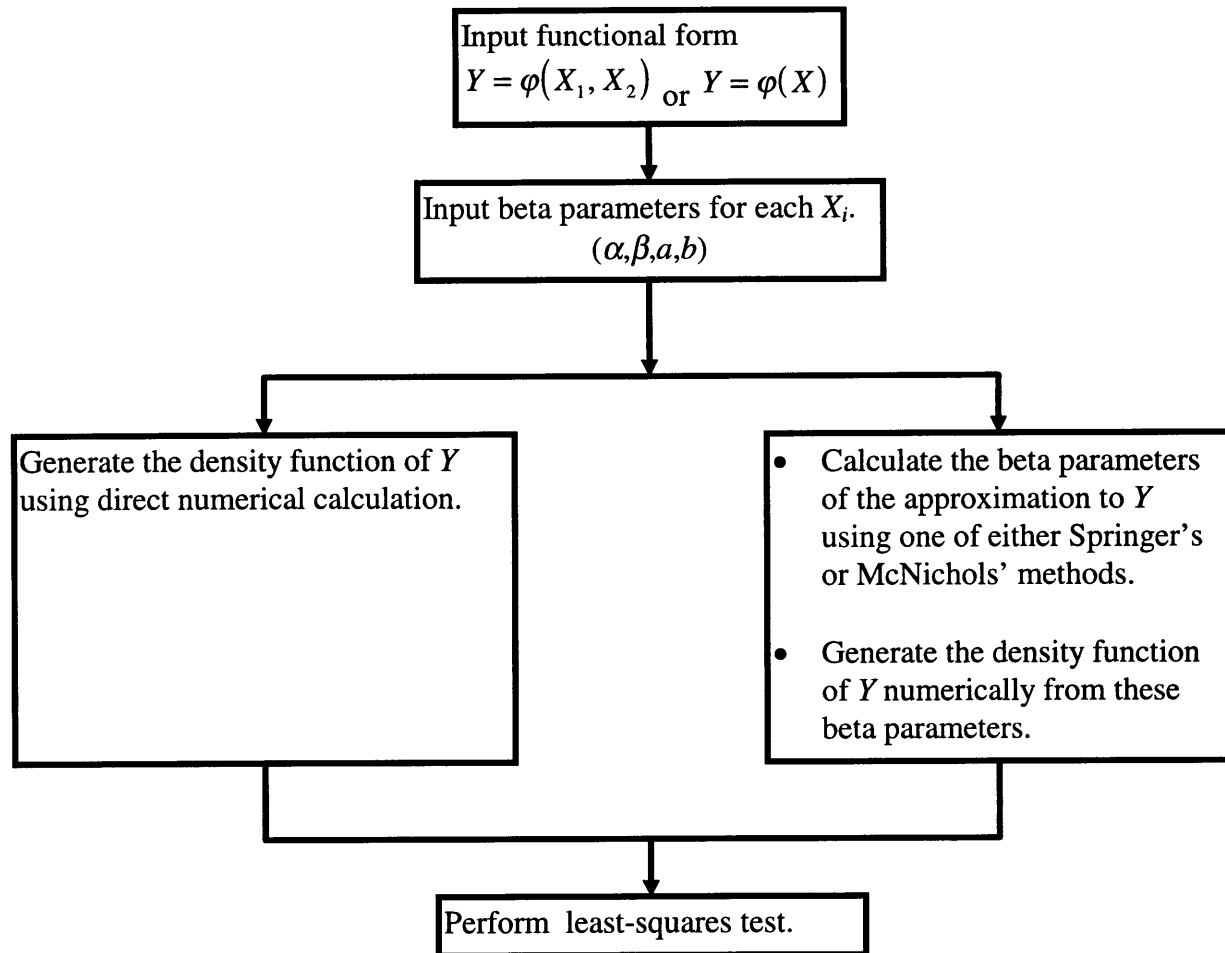


Figure 5.3: A flowgraph of the least-squares routine used to test each of Springer's and McNichols' approximation routines is shown above.

5.3.1 Numerical Convolution of Two Independent Random Variables

The routines for performing numerical Fourier convolution with probability density functions were implemented using 256 point radix-2 Fast Fourier Transforms and Inverse Fast Fourier Transforms. 256 points were necessary in order to yield the desired smoothness of the results. The 128 point vectors f_1 and f_2 were generated using the routine described in Section 5.1. Since Fast Fourier Transforms could be used, execution times were fast (about 30ms), and the resulting density function vectors were smooth for the cases of both addition and subtraction. Since the use of the Fast Fourier Transform to calculate the sum of two probability density functions is a well known technique, the results were not subjected to a chi-squared test.

The routines for performing numerical multiplication and division were implemented with $n = 500$, the minimum number of points necessary to yield the desired smoothness of the results.

The case of multiplication was benchmarked using the input beta density functions $f_1 = f_b(x;2,1,0,1)$ and $f_2 = f_b(x;3,1,0,1)$, since it is known that the product of these two density functions is the beta density function $g = f_b(x;2,2,0,1)$ [4]. Figure 5.4 shows the product of f_1 and f_2 , generated using numerical convolution, superimposed upon g . It can be seen from the plot on the left, that $n = 500$ was sufficiently large to yield a smooth curve, and the plot on the right shows that the numerically generated curve is almost identical to g . Unfortunately, execution times for both the multiplication and division routines were well in excess of 2 seconds. Execution time could have been reduced by implementing the routines in C instead of MATLAB™, but since these routines were written mainly in order to benchmark the analytical approximation routines of Chapter 3, this step was not taken in order to obtain a meaningful comparison.

The accuracies of both the multiplication and division routines were verified using the chi-squared test described in Section 5.2. For the case of $Y = X_1 X_2$, one hundred trials were run with X_1 and X_2 defined on the interval $[0,1]$ and having randomly generated shape parameters such that $1 \leq \alpha, \beta \leq 10$. For the case of $Y = X_1 / X_2$ one hundred trials were run with X_1 and X_2 defined on the interval $[1,2]$ and having randomly generated shape parameters such that $1 \leq \alpha, \beta \leq 10$. All of the approximating approximations generated passed the chi-squared test, showing that both algorithms are reliable.

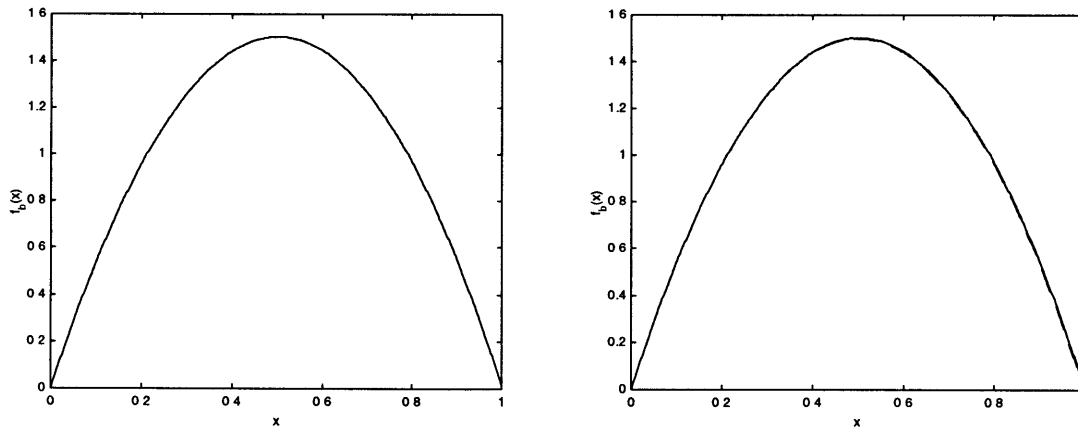


Figure 5.4: Plots of the product of $f_b(x;2,1,0,1)$ and $f_b(x;3,1,0,1)$ generated using numerical convolution (left) and this same curve superimposed upon $f_b(x;2,2,0,1)$ (right) are shown above. The plot on the left shows that the technique yields smooth curves, while the plot on the right shows that accurate curves are yielded through this technique as well.

5.3.2 Numerical Exponentiation of One Random Variable

The routine for performing numerical exponentiation of one random variable to a scalar power directly using Equation (4.26) was implemented with $n = 100$. Execution time was short (30ms) and the resulting curves were sufficiently smooth.

5.3.3 Implementation and Exploration of Monte Carlo Methods

The Monte Carlo routine described in Section 4.3 was implemented for the cases of multiplication, division, and addition of two random variables and exponentiation of one random variable with $p = 10^5$ and $n = 100$. Cheng's [1] algorithm BA was used to generate beta variates from a set of uniformly distributed random numbers (see Section 5.2). Consideration was given to using a significantly more sophisticated algorithm developed by Zechner and Stadlober [17], but once it became clear that the previously described numerical methods were far superior to the Monte Carlo method for our purposes, this idea was abandoned. Execution times were long using the Monte Carlo method and the resulting curves were not smooth. This phenomenon is explained by the fact that the accuracy of the Monte Carlo method used increases only as the square root p [7].

For all of the operations involving two random variables execution times were approximately 1.5 seconds. For the case of exponentiation execution time was about 1 second. Figure 5.5 shows the product of $f_1 = f_b(x;2,1,0,1)$ and $f_2 = f_b(x;3,1,0,1)$, calculated using the Monte Carlo method, plotted with the known result $g = f_b(x;2,2,0,1)$. While the curve generated using the Monte Carlo method follows the desired curve well, it is not nearly as smooth as one would like it to be. Comparison with Figure 5.4 shows that the convolution methods described in the previous section are far superior in terms of the smoothness of their results than the Monte Carlo method described in this section. This can be attributed to the fact that analytic expressions for products, quotients, and sums of two independent random variables X_1 and X_2 contained only one dimensional convolution integrals, which could be readily solved on a computer by sampling the density functions of X_1 and X_2 . For the case of exponentiation of a random variable to a scalar power, the analytic expression to be solved contained no integral at all, rendering its calculation even more straightforward.

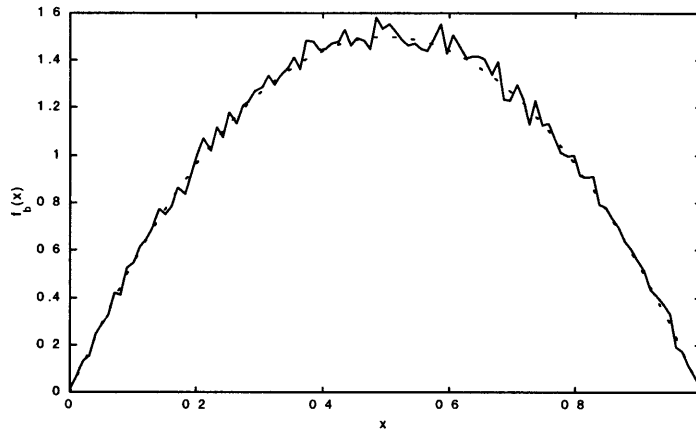


Figure 5.5: The product of $f_1 = f_b(x;2,1,0,1)$ and $f_2 = f_b(x;3,1,0,1)$, calculated using the Monte Carlo method (solid), and the known product $g = f_b(x;2,2,0,1)$ (dashed) are plotted above. Comparison with Figure 5.4 shows that the convolution based methods described in the previous section are far superior in terms of the smoothness of the result than the Monte Carlo method described in this section.

5.4 Implementation and Comparison of Analytical Approximation Methods

In Section 2.2 it was proposed to verify experimentally whether products, sums, quotients, powers, and linear combinations of beta distributions could be well approximated themselves by beta distributions. This section details the implementation and testing of the analytic routines that generate the beta approximations to the products, sums, quotients, powers, and linear combinations of beta distributions. In Section 5.4.1 an algorithm is developed for calculating the beta parameters of a distribution given its central moments. Sections 5.4.2 and 5.4.3 discuss implementation of Springer's and McNichols' methods, described in Chapter 3, and focus specifically on the issue of numerically generating moments of arbitrary order of beta distributions. Implementation of chi-squared and least-squares routines, which are used to test the approximation routines, is the subject of Section 5.4.4. The results of these tests are discussed in Section 5.4.5, and conclusions are drawn about the accuracy of each approximation method for each of the arithmetic operations under study.

5.4.1 Calculation of Beta Parameters from Central Moments

Expressions for calculating the beta parameters of a beta distribution from its first four central moments were given in Section 2.2. It was found experimentally that these expressions sometimes yield negative beta shape parameters (α, β) given certain combinations of the first four central moments. If either of the beta shape parameters is negative, the corresponding beta distribution function will no longer converge to 1 as $x \rightarrow b$, which we know that it must according to Equation (2.2c), implying that negative beta shape parameters are meaningless. The following simple algorithm was therefore developed to transform negative beta shape parameters to positive ones:

1. Calculate α and β using Equation (2.22).
2. If neither α nor β is negative, a valid result has been obtained, and no further steps need be taken. Otherwise let:

$$\alpha' = \alpha, \beta' = \beta, c = \alpha \cdot \beta \quad (5.12)$$

3. If $c > 0$ the new shape parameters are given by $\alpha = |\alpha'|$ and $\beta = |\beta'|$, otherwise the new shape parameters are given by $\alpha = |\beta'|$ and $\beta = |\alpha'|$.

The validity of the above algorithm was verified experimentally using the test methods described in the following sections.

5.4.2 Implementation of Springer's Methods

Approximation of linear combinations and products of two independent random variables using Springer's approximation methods, described in Section 3.1, requires that we calculate only the first four moments (for the case of multiplication) or cumulants (for the case of linear combinations) of each beta variable. Calculation of the first four moments or cumulants of a beta variable may in turn be calculated from each beta variable's central moments. Finally, the central moments may be calculated exactly and readily from the beta shape parameters. Thus, implementation of Springer's approximation method for multiplication, according to Equation (3.18), and linear combinations of independent random variables, according to Equation (3.10), was accomplished without complication.

Implementation of division of two random variables and exponentiation of one random variable to a scalar power was however not as straightforward. The expressions for the moments of the quotient of two independent beta variables and one beta variable raised to a scalar power derived from Springer [16] require moments other than the first four. Specifically, division of two independent random variables requires the negative first, second, third, and fourth moments of the beta variable X_2 , where $Y = \frac{X_1}{X_2}$, and exponentiation of one random variable X to the m^{th} power requires the m^{th} , $2m^{\text{th}}$, $3m^{\text{th}}$, and $4m^{\text{th}}$ moments of X . As no closed form algebraic expressions were available for these moments, they had to be calculated numerically based on Equation (2.3). Numerical integration was accomplished using Romberg's method (see [14]) with a numerical tolerance of 10^{-10} . For this reason, the division and exponentiation algorithms' execution times (about 20ms) are approximately twice those of the methods for multiplication and linear combinations.

5.4.3 Implementation of McNichols' Methods

Implementation of McNichols' first order method for the cases of multiplication, division and exponentiation was straightforward, as these operations require only the first four central moments of each random variable. McNichols' second order method for multiplication also requires only the first four central moments of each random variable. His second order method for the cases of division and exponentiation require both the sixth and eighth central moments of the input random variables. Since no expressions for these central moments were available, they were calculated by evaluating Equation (2.3) using Romberg integration with a numerical tolerance of 10^{-10} . Due to this, run times for McNichols' second order division and exponentiation routines were twice as long (about 20ms) as his other methods.

5.4.4 Description of Test Batteries

Two batteries of tests were run to test the accuracies of each of Springer's and McNichols' methods. One battery of tests was run in order test the accuracy of each method under random variation of the beta shape parameters, α and β , of the input distributions, and a second was run to test each method under random variation of the interval $[a,b]$ upon which each input beta distribution was defined.

Table 5.1 summarizes the first battery of random tests run on Springer’s and McNichols’ approximation algorithms. Each of the beta shape parameters, α and β , for each of the input beta distributions, X_i , was produced with a random number generator. It was chosen to allow the beta shape parameters to vary between 1 and 10, as it is unlikely that very highly skewed distributions would be used to model any real physical or economic phenomenon. For the cases of multiplication and addition each beta distribution is defined on the interval $[0,1]$. For the cases of division and exponentiation each beta distribution is defined on the interval $[1,2]$ in order to avoid ill numerical behavior near $x = 0$. One hundred tests were run for each operation under study, although for the case of exponentiation 4 different scalar powers were applied to each of only 25 randomly generated beta distributions.

Table 5.2 summarizes the second battery of random tests run on Springer’s and McNichols’ approximation algorithms. Each of the beta interval boundaries, a and b , for each of the input beta distributions, X_i , was produced with a random number generator. For the cases of multiplication and addition these were allowed to vary between 0 and 100. For the cases of division and exponentiation, these were allowed to vary between 1 and 100 in order to avoid ill numerical behavior near $x = 0$. α and β of each of the two distributions for each mathematical operation were chosen such that a result was yielded using all three approximation methods that passed a chi-squared test on the interval $[0,1]$. One hundred tests were run for each operation under study, although for the case of exponentiation 4 different scalar powers were applied to each of only 25 randomly generated beta distributions.

Operation	α_1	β_1	a_1	b_1	α_2	β_2	a_2	b_2	# of Tests
$Y = X_1 X_2$	$1 \leq \alpha_1 \leq 10$	$1 \leq \beta_1 \leq 10$	0	1	$1 \leq \alpha_2 \leq 10$	$1 \leq \beta_2 \leq 10$	0	1	100
$Y = X_1 + X_2$	$1 \leq \alpha_1 \leq 10$	$1 \leq \beta_1 \leq 10$	0	1	$1 \leq \alpha_2 \leq 10$	$1 \leq \beta_2 \leq 10$	0	1	100
$Y = \frac{X_1}{X_2}$	$1 \leq \alpha_1 \leq 10$	$1 \leq \beta_1 \leq 10$	1	2	$1 \leq \alpha_2 \leq 10$	$1 \leq \beta_2 \leq 10$	1	2	100
$Y = X^m$	$1 \leq \alpha_1 \leq 10$	$1 \leq \beta_1 \leq 10$	1	2	$m = 0.5, 1.5, 2, 3$			25	

Table 5.1: The beta parameters and number of tests run for the first battery of random tests on Springer’s and McNichols’ approximation algorithms are shown above. All of the beta shape parameters, α and β , were produced using a random number generator.

Operation	α_1	β_1	a_1	b_1	α_2	β_2	a_2	b_2	# of Tests
$Y = X_1 X_2$	9	3	$0 \leq a_1 \leq 100$	$0 \leq b_1 \leq 100$	6	5	$0 \leq a_2 \leq 100$	$0 \leq b_2 \leq 100$	100
$Y = X_1 + X_2$	9	3	$0 \leq a_1 \leq 100$	$0 \leq b_1 \leq 100$	8	6	$0 \leq a_2 \leq 100$	$0 \leq b_2 \leq 100$	100
$Y = \frac{X_1}{X_2}$	2.6	4.6	$1 \leq a_1 \leq 100$	$1 \leq b_1 \leq 100$	9.6	8	$1 \leq a_2 \leq 100$	$1 \leq b_2 \leq 100$	100
$Y = X^m$	4	2	$1 \leq a_1 \leq 100$	$1 \leq b_1 \leq 100$	$m = 0.5, 1.5, 2, 3$			25	

Table 5.2: The beta parameters and number of tests run for the second battery of random tests on Springer's and McNichols' approximation algorithms are shown above. All of the beta interval boundaries, a and b , were produced using a random number generator.

5.4.5 Experimental Comparison of Springer's and McNichols' Methods

The results of the two batteries of tests run on Springer's and McNichols' approximation methods are shown below in Table 5.3 and Table 5.4. The number of tests out of one hundred that failed the chi-squared test at a 1% level of significance for each operation and approximation method is given in the third column of both tables. Whether we chose the chi-squared test or the least-squares test as our performance metric, it is clear that Springer's method is far superior to either McNichols' first or second order methods in terms of accuracy of the approximations it generated. This was especially true for the first battery of tests when the input beta distributions were defined on a much narrower interval than for the second battery of tests. Figure 5.6 compares Springer's and McNichols' approximations for the products of the distributions $f_b(x;4.1,8.15,0,1)$ and $f_b(x;9,9.33,0,1)$. It can be seen that Springer's method performs much better than McNichols' second order method. This result was reflected by the fact that Springer's approximation passed a chi-squared test while McNichols' second order method did not. Although McNichols' multiplication methods were much better in the second battery of tests, his division and exponentiation methods still did not perform nearly as well as Springer's.

Operation	Method	# of Trials that Failed χ^2 Test					Average of Least-Squares Tests
		0.5	1.5	2	3	Total	
$Y = X_1 X_2$	Springer's			0			0.024
	McNichols' 1 st			60			0.360
	McNichols' 2 nd			60			0.320
$Y = X_1 + X_2$	Springer's / McNichols' (see note below)			3			0.0016
$Y = \frac{X_1}{X_2}$	Springer's			8			0.0041
	McNichols' 1 st			32			0.011
	McNichols' 2 nd			27			0.017
$Y = X^m$				<i>m</i>			
	Springer's	0	0	0	0	0	0.007
	McNichols' 1 st	0	1	2	12	15	0.056
	McNichols' 2 nd	1	0	3	10	14	0.043

Table 5.3: The results of the first battery of tests run on Springer's and McNichols' approximation methods are shown above (see Table 5.1). For each arithmetic operation and approximation method, the number of trials out of one hundred that failed a chi-squared test is given, along with the average over all one hundred trials of the least-squares test (see Equation (5.11)). Note that McNichols' first and second order methods are identical to Springer's method for the case of addition (see Section 3.2).

Operation	Method	# of Trials that Failed χ^2 Test					Average of Least-Squares Tests
$Y = X_1 X_2$	Springer's	0					4.8×10^{-6}
	McNichols' 1 st	0					4.8×10^{-6}
	McNichols' 2 nd	0					4.8×10^{-6}
$Y = X_1 + X_2$	Springer's / McNichols' (see note below)	0					0.0060
$Y = \frac{X_1}{X_2}$	Springer's	3					0.010
	McNichols' 1 st	22					0.048
	McNichols' 2 nd	18					0.070
$Y = X^m$		<i>m</i>					
		0.5	1.5	2	3	Total	
	Springer's	0	0	0	0	0	1.9×10^{-6}
	McNichols' 1 st	0	1	6	13	20	1.9×10^{-4}
McNichols' 2 nd	1	0	6	14	21	2.1×10^{-4}	

Table 5.4: The results of the second battery of tests run on Springer's and McNichols' approximation methods are shown above (see Table 5.2). For each arithmetic operation and approximation method, the number of trials out of one hundred that failed a chi-squared test is given, along with the average over all one hundred trials of the least-squares test (see Equation (5.11)). Note that McNichols' first and second order methods are identical to Springer's method for the case of addition (see Section 3.2).

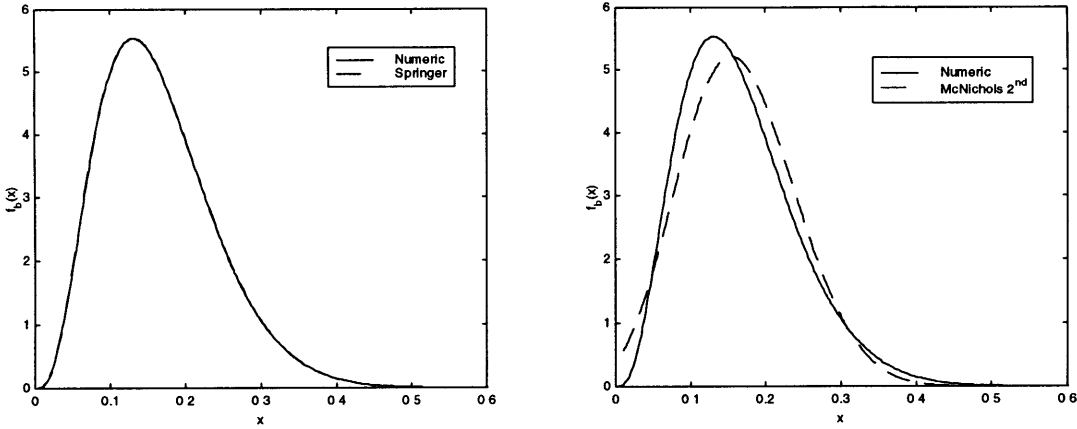


Figure 5.6: Plots of the approximation to the product of $f_b(x; 4.1, 8.15, 0, 1)$ and $f_b(x; 9.933, 0, 1)$ shown on the interval $[0, 0.6]$ are shown above for Springer's method (left) and McNichols second order method (right). It can be seen that Springer's method yields the better approximation.

A plot of the one product trial using Springer's method that passed a chi-squared test but yielded a high least-squares value of 2.06 is shown in Figure 5.7. The approximation generated numerically is highly skewed, and therefore does not lend itself well to being approximated by a beta distribution. Although Springer's approximation is excellent on the interval $[0.02, 1]$, most of the probability of the desired result as well as the approximation is concentrated in the region $[0, 0.02]$ where the approximation is poor. Thus when Springer's product method yields a density function that approaches infinity at either its right or left hand limit, ($\Rightarrow \alpha < 1$ OR $\beta < 1$), the results should be examined against a numerically generated benchmark before they are trusted. Otherwise, Springer's multiplication method is highly reliable. Note that Springer's method always yields the best beta approximation to the true result, since the first four moments of Springer's approximation exactly match those of the true result (see Section 2.1.3). Thus, when Springer's method yields a poor approximation to a given density function, we may not use a beta density function to approximate this density function.

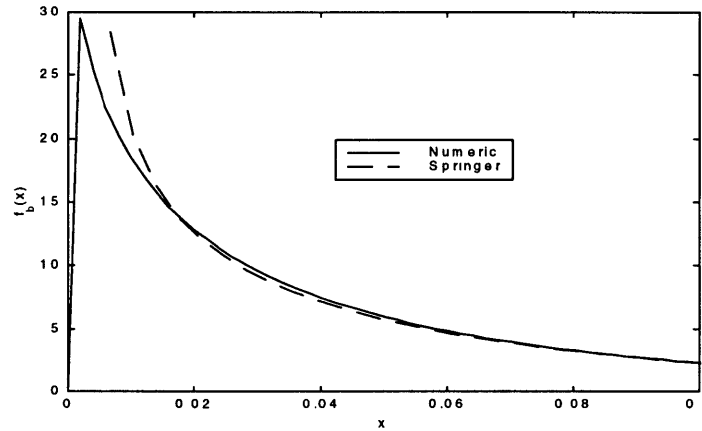


Figure 5.7: Plots of Springer’s approximation to the product of $f_b(x;1.24,3.81,0,1)$ and $f_b(x;1.12,4.46,0,1)$ and the numerically generated benchmark density function on the interval $[0,0.1]$ are shown above. In this case, Springer’s method yields a result of $f_b(x;0.57,21.70,0.00,1.79)$ and a least-squares metric of 2.06.

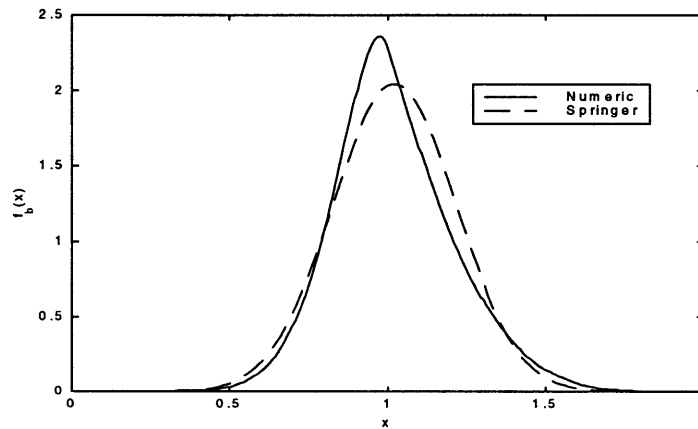


Figure 5.8: Plots of Springer’s approximation to the sum of $f_b(x;9.25,2.11,0,1)$ and $f_b(x;1.12,4.33,0,1)$ and the numerically generated benchmark density function on the interval $[0,2]$ are shown above. In this case, Springer’s method yields a result of $f_b(x;49.75,49.75,-0.92,2.96)$ and a least-squares metric of 0.0177.

For the case of addition of two random variables, most of the trials passed a chi-squared test as well. In contrast to the case of multiplication, approximations to sums of two random variables tended to be worse when the resulting distribution was fairly sharp. A plot of an approximation that failed the chi-squared test is shown in Figure 5.8. When Springer’s addition method yields a density function that is sharp, i.e. α and β are both large, the results should be examined against a numerically generated benchmark before they are trusted.

Although Springer's method was also the best overall for approximating the quotient of two beta distributions, it was found that in a few isolated cases, McNichols' first and second order methods yielded slightly better approximations. By the same token, McNichols' first order method tended to yield better results than the second order method. Since both Springer's method and McNichols' second order method both required the use of numerical Romberg integration, it is believed that small numerical inaccuracies were the cause of the unexpected results. Figure 5.9 however shows that when McNichols' first order method is best, it is not much better than either his own second order method or Springer's method.

Many quotients did not lend themselves well to approximation by a beta distribution. Figure 5.10 shows one of the worst approximations using Springer's method. Note that the desired density function is fairly sharp, and therefore does not allow itself to be approximated well by a beta density function. Unfortunately, it is difficult at present to predict when Springer's method will yield an unsatisfactory approximation for the case of division.

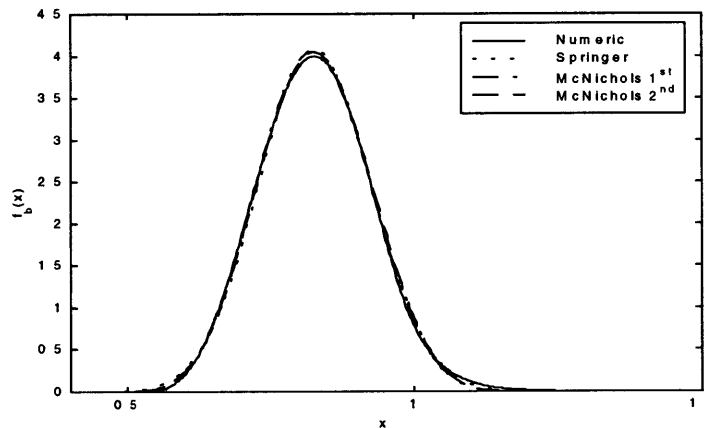


Figure 5.9: Plots of the approximation to the quotient of $f_b(x;4.74,3.75,1,2)$ and $f_b(x;8.87,1.36,1,2)$ on the interval $[0.4,1.5]$ are shown above for the three methods tested. Both McNichols' first and second order approximations are more accurate than Springer's, but it is difficult to discern that from this figure.

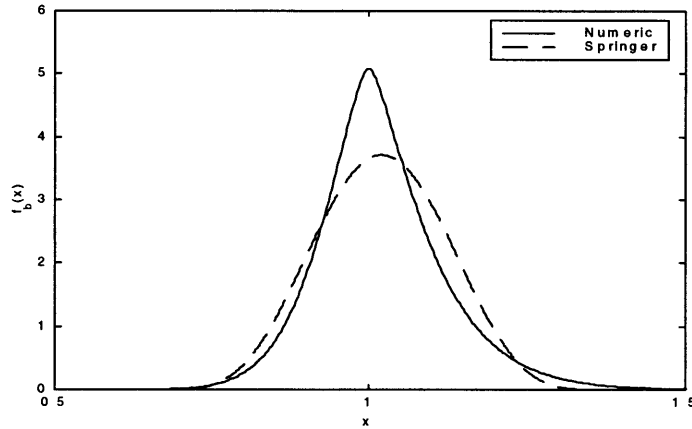


Figure 5.10: Plots of Springer's approximation to the quotient of $f_b(x;8.15,1.53,1,2)$ and $f_b(x;6.43,1.45,1,2)$ and the numerically generated benchmark density function on the interval $[0.5,1.5]$ are shown above. In this case, Springer's method yields a result of $f_b(x;6.10,6.10,0.65,1.39)$ and a least-squares value of 0.2151.

For the case of exponentiation of one random variable to a scalar power m using Springer's method, all of the trials passed a chi-squared test. It was however found that when m was less than 0 or greater than 3, Springer's method became unreliable. Both McNichols' first and second order methods performed worse than Springer's for all m tested.

The conclusion to be drawn from both batteries of tests is that Springer's approximation methods outperform McNichols' methods in terms of accuracy in almost all situations for the arithmetic operations of multiplication, addition, division, and exponentiation to a scalar power m , such that $0 < m \leq 3$. For the case of addition, Springer's method seems to generate very accurate results as long as the beta shape parameters of the input distributions are between 1 and 10. For the cases of multiplication and exponentiation, most of the approximations generated using Springer's method are accurate. When results are generated with shape parameters less than 1, caution should be used in accepting these results. Approximation of the quotient of two beta distributions does have the potential to lead to inaccurate and unexpected results. One topic of future research could therefore be to develop a new method to more accurately approximate the quotient of two beta variates or at least predict when approximations using Springer's method will be inaccurate.

5.5 Comparison of Execution Times

Table 5.5 shows execution times in seconds of the direct numerical and Monte Carlo methods as well as Springer's and McNichols' analytical approximation methods. For the case of addition, the convolution method described in Section 4.1.1 was much faster than the Monte Carlo method described in Section 4.3, since the former method could be implemented using a radix-2 Fast Fourier Transform algorithm. Calculation of the density function of one random variable raised to a scalar power was significantly faster by direct calculation according to Equation (4.26) than with the Monte Carlo method. When performing multiplication and division, the Monte Carlo method proved considerably faster than the convolution methods described in Section 4.1.1. However, the curves resulting from Monte Carlo calculations proved inferior in terms of their smoothness in comparison to the other methods. This problem could be remedied by increasing the number of points sampled from each input distribution, but as the accuracy of the Monte Carlo method increases only as the square-root of the number of points sampled [7], large gains in accuracy would drastically increase execution times.

Springer's and McNichols' analytical approximation methods proved much faster than direct numerical calculation or calculation using a Monte Carlo approach. Note that Springer's and McNichols' second order routines for division and exponentiation were slower than the other analytical approximation routines due to their use of Romberg integration .

Method	Operation	Multiplication	Division	Addition	Exponentiation
Direct		2.37	2.7	0.03	0.03
Monte Carlo		1.51	1.4	1.39	0.91
Springer		0.01	0.02	0.01	0.02
McNichols 1 st		0.01	0.01	0.01	0.01
McNichols 2 nd		0.01	0.02	0.01	0.02

Table 5.5: Execution times in seconds of all of the numerical and analytical methods explored for generating approximations to products, quotients, sums, and powers of independent beta random variables are shown above.

6 Conclusions and Further Research

The overall objective of this thesis was to investigate methods of approximating probability densities of functions of independent random variables for the purpose of using in the MAESTrO software tool. Stated mathematically, if X_1, X_2, \dots, X_n are independent real random variables with density functions $f_1(x), f_2(x), \dots, f_n(x)$ and $Y = \varphi(X_1, X_2, \dots, X_n)$, where φ is any arbitrary real function, reasonable approximations were found to the probability density function of the random variable Y , $g(y)$, using the density functions of X_1, X_2, \dots, X_n .

While it is theoretically possible in many instances to find a closed form analytical expression for $g(y)$, this expression can be very complicated, difficult to derive, and impossible to evaluate on a computer. It was therefore desirable to find a reasonably good approximation to $g(y)$ using simple numerical algorithms that require a fraction of the computational time and resources of their analytical counterpart. This thesis compared methods for approximating the density of the function of independent random variables.

The overall objective of this thesis was to investigate the following three methods for performing calculations with random variables and to compare them with regard to their relative speed and accuracy:

1. Direct numerical calculation of $g(y)$ based on analytical expressions
2. Monte Carlo approximation to $g(y)$
3. Analytical approximation to $g(y)$ using the moments or cumulants of X_1, X_2, \dots, X_n

Experimental results showed that approximating $g(y)$ numerically, based on analytical expressions, was the most accurate for the cases of multiplication, division, and addition of two random variables and exponentiation of one random variable to a scalar power. For the cases of addition and exponentiation, these numerical approximations could also be generated quickly. Approximating $g(y)$ using a Monte Carlo approach was slow as well as inaccurate in comparison, due to the fact that fairly simple analytic expressions exist for the arithmetic operations under study.

The direct numerical method however needs to be further refined so that it can be integrated into a software package for performing stochastic system analysis and evaluation. In order for this

method to be used for practical stochastic analysis of complex systems, an algorithm must be developed that can rapidly evaluate the Mellin convolution integrals of Section 4.1.1. Compression algorithms also need to be developed or existing algorithms utilized in order that storage of many stochastic variables does not consume an excessively large amount of computer memory resources.

Springer's analytical approximation method was clearly better than either of McNichols' methods at performing the arithmetic operations under study. For shape parameters of the input beta distributions greater than 1 but less than 10, Springer's method was found to reliably generate accurate approximations to products and sums of two independent random beta variates, as well as to one random beta variate raised to a scalar power between 0 and 3. Springer's approximation methods also require significantly less computer memory than direct numerical methods, as each result is described uniquely and completely by four parameters, and less computation time, due to their simplicity. Unfortunately, Springer's method was found to be a bit unreliable when calculating the quotient of two independent random variables.

Despite their problems, Springer's methods are ready to be integrated into MAESTrO. In order to avoid erroneous results, the accuracy of calculations that meet the following criteria should be verified using a numerically generated benchmark:

- All division operations
- Exponentiation of a beta distribution to a scalar power m that is less than 0 or greater than 3
- Multiplication operations that generate highly skewed beta distributions
- Addition operations that generate sharp beta distributions
- Calculations involving beta distributions with shape parameters less than 1 or greater than 10

Although it is unlikely that sharp or highly skewed beta distributions would arise as input models to any systems, they can come about as the result of previous calculations performed on beta distributions with more symmetric and nearly uniform distribution shapes. The Central Limit Theorem tells us that as we sum more and more independent random variables, the result approaches a Gaussian distribution. When we attempt to approximate something close to a Gaussian with a beta distribution, the interval upon which the beta distribution is defined, $[a,b]$, and its shape parameters, α and β , become large, due to the fact that a Gaussian has infinitely long tails. Multiplication of many independent random variables is not as well understood, but appears to yield results that become more and more skewed. This phenomenon translates to beta distributions with small shape parameters, α and β . Performing calculations with very sharp distributions becomes difficult due to the floating point limits imposed by computers (see Section

5.1). Performing calculations with highly skewed distributions is also difficult, as these distributions may asymptotically approach infinity at either their right or left hand limits.

In order that Springer's method can be more confidently incorporated into software, it is desirable to find more symmetric and nearly uniform beta approximations to sharp or highly skewed beta distributions. In the case of sharp beta distributions, almost all of the probability density is concentrated at the center of the interval $[a,b]$. We would like to find an approximation to this sharp distribution that has much smaller shape parameters and that is defined over a much smaller interval than the original sharp distribution. In the case of highly skewed beta distributions, we would like to find an approximation that is defined over a larger interval and with larger shape parameters than the original skewed distribution. Finding algorithms to perform these mappings would greatly assist practical implementation of Springer's methods into software, and provides a potentially interesting topic for further research.

Bibliography

- [1] Cheng, R. C. H., "Generating Beta Variates with Nonintegral Shape Parameters," *Communications of the ACM*, Vol. 21, No. 4, pp 317-322, 1978.
- [2] S. Y. Dennis, "On the Distribution of Products of Independent Beta Variates," *Communications in Statistics - Theory and Methods*, Vol. 23, No. 12, pp 1895-1913, 1994.
- [3] W. J. Fabrycky, B. S. Blanchard, *Life-Cycle Cost and Economic Analysis*. Englewood Cliffs: Prentice Hall, 1991, Chapter 1.
- [4] D. Y. Fan, "The Distribution of the Product of Independent Beta Variables," *Communications in Statistics - Theory and Methods*, Vol. 20, No. 12, pp. 4043-4052, 1991.
- [5] B. Gnedenko, *The Theory of Probability*. Moscow: Mir Publishers, 1975, Chapter 4.
- [6] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*. New York: John Wiley and Sons, Inc., 1996, Chapter 1.
- [7] J. M. Hammersly and D. C. Handscomb, *Monte Carlo Methods*. London: Chapman and Hall, 1965.
- [8] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*. New York: John Wiley & Sons, 1995, Chapter 25.
- [9] M. Kendall, A. Stuart, and J. K. Ord, *Kendall's Advanced Theory of Statistics*. 5th ed. New York: Oxford University Press, 1987.
- [10] The MathWorks, Inc., Natick, MA. *Using MATLAB*, 1996.

- [11] G. R. McNichols, *On the Treatment of Uncertainty in Parametric Costing*. Doctoral Thesis, School of Engineering and Applied Science, George Washington University, 1976.
- [12] T. Neff, *Ansätze zur Verarbeitung unvollständiger und unsicherer Informationen im Rahmen der technisch-wirtschaftlichen Produktbewertung*. Internal Technical Report, Daimler-Benz AG, 1997.
- [13] V. S. Pugachev. *Probability Theory and Mathematical Statistics for Engineers*. Oxford: Pergamon Press, 1984.
- [14] W. H. Press, et al., *Numerical Recipes in C: The Art of Scientific Computing*. 2nd ed. Cambridge: Cambridge University Press, 1992.
- [15] T. T. Soong, *Probabilistic Modeling and Analysis in Science and Engineering*. New York: John Wiley & Sons, 1981.
- [16] M. D. Springer, *The Algebra of Random Variables*. New York: John Wiley & Sons, 1979.
- [17] H. Zechner and E. Stadlober, "Generating Beta Variates via Patchwork Rejection," *Computing*, Vol. 50, No. 1, pp 1-18, 1993.
- [18] M. Iqbal, "On Regularized Numerical Inversion of Mellin Transform," *Computer Physics Communications*, Vol. 90, pp 293-300, 1995.
- [19] K. K. Afridi, *A Methodology for the Design and Evaluation of Advanced Automotive Electrical Power Systems*. Doctoral Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1998.
- [20] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*. New York: McGraw-Hill, 1969.