International Journal of Modern Mathematical Sciences, 2014, 10(2): 189-198

Modern Scientific Press

International Journal of Modern Mathematical Sciences ISSN: 2166-286X

Journal homepage: www.ModernScientificPress.com/Journals/ijmms.aspx *Florida, USA*

Article

Application of Computerized Monte Carlo Simulation in Evaluating Definite Integrals and Testing of the Properties of Probability Density Functions

Ogwumu, David. O[1], James Friday. E[2] and Abubakar S[3]

[1,2]*Department of Mathematics and Statistics, Federal University Wukari, P.M.B 1020, Wukari – Taraba State, Nigeria*

[3]*Department of Mathematics and Statistics, Federal University of Technology, Minna, P.M.B 65, Minna – Niger State, Nigeria*

*Author to whom correspondence should be addressed: onahdavid2010@gmail.com (David Onah), chatwitfrayoo@yahoo.com (Friday James).

Article history: Received 27 August 2013, Received in revised form 20 November 2013, Accepted 16 January 2014, Published 23 May 2014.

Abstract: This work explores a certain application of the Monte Carlo simulation technique in evaluating definite Integrals and x-rays the complexity of solving realistic models by mathematical methods to arrive at an analytic solution. The study has shown that certain complex problems that cannot be solved analytically could be subject to simulations to provide approximate solutions. The work equally studies numerous definite Integrals but provides a result for few through which it is possible to make a generalization in the end of the study. Five probability distributions were highlighted in this work (among others that were investigated) and their properties by Monte Carlo simulation using a PASCAL program where random numbers were generated after numerous trials. Some areas of applications of simulation and the probability distributions studied have been discussed in this study alike.

Keywords: Monte Carlo Simulation, undeterministic, probability, Definite Integrals, Area under curve, exact solution, approximate solution, random variables, parametized.

Mathematics Subject Classifications (2000): Probability Application, Monte Carlo Simulation, Probability Density Functions

1. Introduction

This work is born out of the fact that, there are certain classes of Integrals/complex Integrals/ multidimensional integrals /area under curves that have proven difficult to be evaluated analytically. Meanwhile, the Monte Carlo simulation technique is particularly useful in applications that include stochastic (*undeterministic*, *unpredictive* and time dependent) elements and also in estimating values that cannot be mathematically deduced [2]. More broadly, Monte Carlo methods are useful for modeling phenomena with significant uncertainty in inputs [3]. A classic use is for the evaluation of definite integrals, particularly multidimensional integrals with complicated boundaries. The method also involves using random numbers and probability to solve problems.

An application of simulation methods to the estimation of difficult integrals began with the work of [8] and others working in nuclear physics in the 1940s [9]. The basis for Monte Carlo technique is the strong law of large numbers which states that a series of independent and identically distributed random variables, normalized by the number of terms in the series, will converge to the mean of any of the terms in the series, and that enables the evaluation of integrals through simulation[7].

The Monte Carlo name is taken from the Monte Carlo, the capital city of Monaco which was a center for gambling [10]. The Monte Carlo method is a numerical technique for calculating the integrals by using a sequence of random numbers. That is, it consists of algorithms for the approximate evaluation of complex definite integrals, usually multidimensional ones [10].

Certain probability density functions cannot be evaluated analytically and calls for the use of simulation as it is useful when analytic methods have failed. The Monte Carlo method of simulation is therefore employed in this research to test the properties of some probability density functions.

The evaluation of probability functions can be done manually or with the help of desk calculators. However, when the function is too complex and in order to obtain sufficiently precise results, an electronic computer becomes an indispensable tool in such evaluations. This is also confirmed by the remark of a scholar [11] that in the study of computer systems modelling and simulation we need a proper knowledge of both the techniques of simulation modelling and the simulated systems themselves. Therefore, the reliance of the Monte Carlo technique on repeated computation and random numbers makes the need for a computer program inevitable. This sequence of unpredictable numbers is used to represent unpredictable features of a system and is assigned to possible outcomes. Hence, assigning range of numbers to different possible outcomes then using numbers to decide which of the outcomes occur describes the Monte 'Carlo technique of simulation.

1.1. Statement of the Problem

In most cases where the probability density function of a random variable, f(x) is too complex for integral evaluation and to obtain results of high degree of precision, some approximation techniques for testing the properties are used. The Monte Carlo method, which is one of such techniques, is not out of place in this regard to test the properties of certain probability density functions with the aid of an electronic computer.

1.2. Scope of the Study

This research work covers aspects of probability theory and simulation with the use of an electric computer. The probability density functions of some distributions defined in terms of continuous random variable and their applications are studied. The core of the work centers on applying the Monte Carlo method of simulation to test the properties of these probability distributions defined in the statement of the problem. Hence, the area of coverage of this work is limited to probability theory and simulation.

1.3. Aim and Objective of the Study

This research project is aimed at using a Monte Carlo simulation technique in evaluating probability density function f(x) of continuous probability distributions. The study also aims at examining some of these distributions and their areas of relevant applications.

The objective and set goal for this work is to deduce that the integral value of the density functions under study tends to unity. This will provide sufficient reasons or evidence why the Monte Carlo method is very useful when it is infeasible or impossible to compute an exact result. The study is therefore carried out in order to arrive at a dependable conclusion that simulation is useful when analytical methods have proved futile.

1.4. Research Hypothesis

The Monte Carlo technique relies on the use of random numbers and this makes it suitable for computation by the computer. However, not all random numbers are consistent for use in carrying out such experiment. There is therefore need to test for uniformity and independence to test whether or not these numbers generated are random or not. The following hypothesis would be tested in the course of the work.

Kolmogorov Smirnov test for uniformity

H_o : The sequence of numbers is uniform

H_1: The sequence of numbers is not uniform

Runs test for independence

H_o: The sequence of numbers is independent

H_1: The sequence of numbers is dependent

1.5. Probability Distribution Study

Wood [1] Described probability distributions in terms of their probability density functions, occurrence and areas of application. His view of probability description is what is studied in this work. These distributions were considered in this research as follows:

Uniform Distribution

In probability theory and statistics, the continuous uniform distribution is a family of probability distributions such that for each member of the family, all intervals of the same length on the distribution's support are equally possible. The support is defined by two parameters a and b, and the distribution is abbreviated by U(a, b). The probability density function of the uniform distribution is:

$$f(x) = \frac{1}{b-a}, a \leq x \leq b$$

The values of *a* and *b* are usually unimportant because they do not alter the value of the integral f(x). The distribution is used in the generation of random variates that are in turn used in simulation processes.

Exponential Distribution

The exponential distribution is a class of continuous probability distribution. It describes the times between events in a process in which events occur continuously and independently at a constant average rate. The density of the distribution is:

$$f(x; \lambda) = \lambda e^{-\lambda x}$$

where $\lambda>0$ is a parameter of the distribution called the rate parameter. The distribution occurs naturally when describing the lengths of the inter-arrival times in a homogenous process. It has a wide range of applications from reliability theory, study of radioactive particle decay to study of queuing theory where service times of agents of a system are modelled as exponential distribution.

Normal Distribution

Many of the techniques used in applied statistics are based on the normal distribution. A random variable is defined to be normally distributed if its function is given by;

$f(x; \mu, \sigma) = \frac{e^{-\left(\frac{x-\mu}{2\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$, $-\infty<\mu<\infty$ and $\sigma>0$

The normal distribution is used in carrying out statistical test and estimation. It is often applied in physical phenomena such as light intensity and also in financial analysis such as changes in exchange rates, prices and stock market indices.

Beta Distribution

This is a family of continuous probability distributions defined on the interval [0,1] and *parametized* by two shape parameters α and β. The density function for the beta distribution is:

$$f(x;\alpha,\beta)=\frac{\Gamma(\alpha+\beta)x^{\alpha-1}}{\Gamma(\alpha)\Gamma(\beta)}(1-x)^{\beta-1}, \qquad \alpha,\beta>0$$

Beta distributions are used extensively to model events which are constrained to take place within an interval defined by minimum and maximum value. It is used along with other distributions in critical path method and other control systems to describe time to completion of a task.

Gamma Distribution

A continuous random variable is said to have a gamma distribution with parameters α and β if the probability density function is given by:

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)}, \qquad \alpha,\beta>0$$

The gamma distribution has special applications in reliability testing

2. Materials and Method

Suppose the function *f*, describes the probability density in terms of the input variable *x* of a probability distribution. Assume that the function f takes only non-negative values and that it is bound in the interval [a, b] by *m*>0, then the graph of f over [a, b] is entirely contained in a rectangle of dimensions [b, a] by m as shown in figure 1 [6].

$$A=\int_a^b f(x)dx \tag{i}$$

The value of the probability density function is the measure of the shaded area, A in the rectangle. The relative area is

$$p=\frac{A}{(b-a)m} \tag{ii}$$

The relative area is the number between 0 and 1. This number can be interpreted as a probability. If a point is picked at random from the points of the rectangular area, then the probability that the point lies in the shaded area under the curve is the value p in the equation (ii)

The value $A=\int_a^b f(x)dx=p(b-a)m$

The Monte Carlo method provides a way of obtaining the probability, p, directly by the idea of relative frequency in probability. The method applied here involves performing an experiment of choosing a point at random from the rectangular plane and noting whether or not the point lies in the shaded area, A. The experiment is repeated large number of times and p is approximately the frequency of selecting points in the area A [6]. That is, the proportion of points in the area is obtained by

$$p = \frac{numberoftimespointliesinA}{numberoftrials}$$

From the theory of probability, p improves as the number of trials increases. The experiment, conducted at certain number of trials as programmed by the computer gives an accurate estimate of *p* and hence an accurate estimate of $A = \int_a^b f(x)dx$

A point can be chosen at random from the rectangular region by first choosing its *x*-coordinate at random and then its *y*-coordinate at random.

The *x*-coordinate must lie between the numbers a and b and the *y*-coordinate between 0 and m. A computer program is designed in PASCAL (Turbo Version 1.5) to generate random number between 0 and 1 with an *equiprobable* distribution. Using these random numbers, a point in the desired region with coordinates (x_o, y_o) can be determined, where

$$x_0 = a + rand * (b - a)$$

$$y_o = m * rand$$

Where *a*, *b* and *m* are arbitrary values chosen based on the probability distribution under consideration and *rand* the random number

To decide if the randomly chosen point (x_o, y_o) belongs to A, we compute $f(x_o)$. That is, substitute the value of xo into the given probability density function and determine whether the inequality

$y_0 < f(x_o)$ (iii)

The above is valid. If it is, then the point belongs to A, otherwise it is not.

Hence, to obtain our approximate value for p, we generate N points in the rectangular region in the manner described and keep track of what proportion of times the coordinates of the points satisfy the inequality in equation (iii)

Test for uniformity and independence of random numbers generated

As stipulated in the statement of the research hypothesis, not all numbers are consistently random for use in carrying out a research that could achieve the set goal as this. The Runs test and the one-sample Kolmogorov Smirnov test at 5% level of significance are used to test for independence and uniformity respectively of the numbers generated by the computer program. A sample of 30 numbers generated is used for the test. The test is performed using Statistical Package for Social Sciences (SPSS) and the results and conclusions are shown in appendix II

3. Results and Discussion

The computer program designed to carry out the Monte Carlo simulation is displayed on appendix VII. Five probability distributions were studied and the table below shows their respective total area under curve as tested by the Monte Carlo simulation.

The total area under curve gives the area of the shaded portion of figure 1. That is, the integral $A = \int_a^b f(x)dx$ and this tends to unity. Also the value of this total area increases as the number of trials is increased per distribution.

Moreover from table 1 below, all the distributions we evaluated using our method of study, their total area under curve gives values that are approximately one (1) which conforms to reality since the total area under the curve of any distribution must not exceed unitary. This fact thus recommends our method as one of the best simulation methods of evaluating complex functions of distributions which may be difficult to be solved analytically.

This research work is being carried out to arrive at a reasonable conclusion that the Monte Carlo technique of simulation is fit for Investigating the reality of evaluating complex definite Integrals, which would provide sufficient reasons to apply the Monte Carlo method where an exact or analytical solution cannot be achieved since its approximate solution provided is very close to the exact solution.

Table 1: Result of the Computerized Monte Carlo Simulation for Evaluating some Definite Integrals.

Distribution	**Number of trials**	**Total Area Under Curve in the Interval [a, b]**	**Exact Solution For Total Area Under The Curve f(x)**	**Absolute Error**
Continuous uniform distribution a=1, b=11, m=0.2, interval [0, 10]	3500 400	0.96343 0.96700	1 unit 1 unit	0.03657 0.03300=3.3%
Normal distribution a=0, b=2.36, m=6.20	2000 2500	0.95840 0.99498	1 unit 1 unit	0.04160 0.00502=0.5%
Exponential distribution a=0, b= 1.36, m=7.2, λ=0.85	1000 1500	0.97561 0.99552	1 unit 1 unit	0.02439 0.00448=0.5%
Gamma distribution a=2, b=11, m=0.2, α=3, β=2	500 700	0.98100 0.99964	1 unit 1 unit	0.01900 0.00036=0.4%
Beta distribution a= 0.15, b= 0.99, m=2.2, α=2,β=3	1000 1500	0.95542 0.95726	1 unit 1 unit	0.04454 0.04274=4.2%

Source: Result of the Computerized Monte Carlo Simulation for Evaluating some Definite Integrals verified by the author.

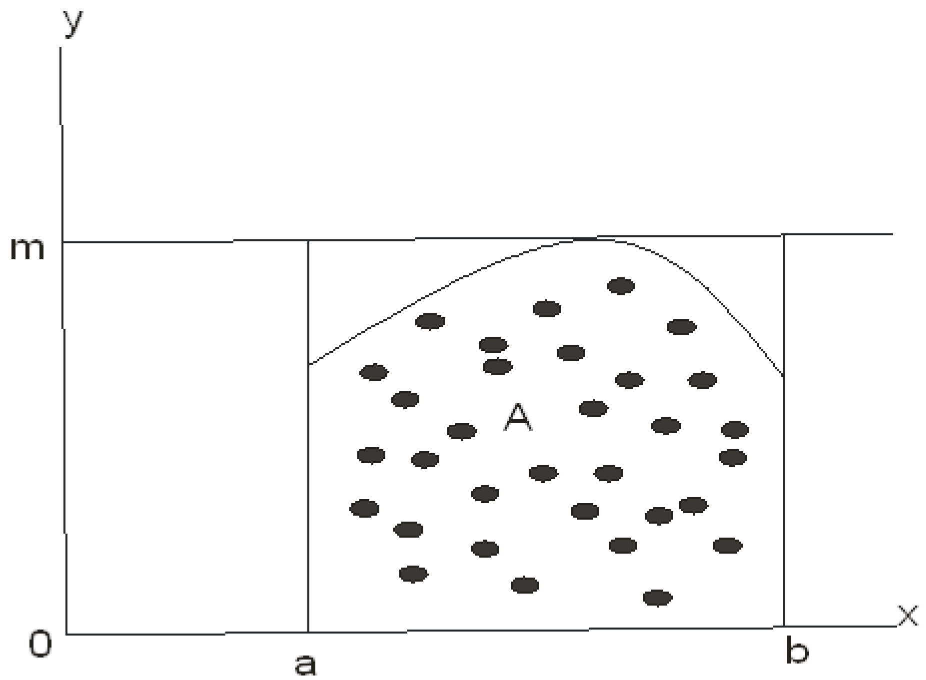

Figure 1: Integral Area under the curve/ The graph of the shaded area A, as the value of the density function f(x) covers the interval [a, b].

4. Conclusion

It is therefore sufficiently observed from the computerized experiment performed and the output obtained as tabulated in table 4.1 that, the Monte Carlo technique is considered as an alternative approach to provide an approximate solution where exact solution cannot be arrived at when dealing with complex definite Integrals. This is consequent upon obtaining an average negligible percentage error less than 5% placing our method at about 95% of accuracy (in terms of absolute deviation from exact solution) as deduced from table 4.1 above. It was observed from the table 4.1 that, the higher the number of trials, the closer the rate of convergence to the exact solution. Hence, it is therefore recommended for use in solving complex problems whose solutions are difficult to be met by other mathematical or statistical methods and approaches.

References

[1] Alexander M. Wood, Franklin A. Graybill, Duane C. Boes; *Introduction to theory of Statistics*, McGraw Hill Int'l Company, Tokyo, **1974**.

[2] Arsham H., Performance Extrapolation in Discrete-event Systems Simulation, *Journal of Systems Science*, 27(9) **(1996)**: 863-869

[3] Averill M. Law, W. David Kelton, *Simulation Modelling and Analysis,* McGraw Hill Inc, U.S.A., **1991.**

[4] Berg, Bernd A. *Makov Chain Monte Carlo Simulations and their Statistical Analysis (With Web based FORTRAN code)* ISBN 981-238-935-0, **2013**.

[5] Catlow B., *Quantitative Method for computing students, An interactive learning approach,* Guansey press 1st ed, Great Britain, **1993.**

[6] Elwood S. Buffa, Rakesh K. Sarin, *Modern Production and Operations Analysis,* McGraw Hill/Irwin, New York, **1994**.

[7] Graham V. Weinberg, Ross Kyprianou, Approximation of Integrals via Monte Carlo Methods with an application to calculating Radar Detection Probabilities, DSTO, South Australia, Australia, **2005**.

[8] Jerome Sacks, William J. Welch, Toby J. Mitchell, Hency P. Wyon, Design and Analysis of computer experiments, Statistical Science*, a review journal of the Institute of Mathematical Statistics,* 4(4) **(1989):** 409-435

[9] Kahn. H, *Random Sampling (Monte Carlo) technique in Neutron Attenuation Problems,* I-Nucleonic, pp 27-37, **1950**.

[10] Karrar D.H Mohammed, *Monte Carlo Numerical Methods, a lecture @* www.mathcomp.uokufa.edu.iq, retrieved on 15th February, **2014**.

[11] Korve K.N, Chukwudolue C. N, *The evaluation of definite integrals using the Monte Carlo Method, Nigerian Journal of Mathematics and applications,* 9(**1996**): 239-249

[12] Owen L. Davis, Peter L. Goldsmith, *Statistical methods in Research and Production,* John Wiley and Sons Inc., New York, **1988.**

[13] Robert A. Stern, Nancy Stern, *Introduction to Computers and Information Processing,* John Wiley and Son Inc., Canada, **1979.**

[14] Steven Nahmias, *Production and Operations Analysis*, McGraw Hill/Irwin, New York. **2001**

[15] Udom, A. U, *Essentials of Statistics, Magut Business Enterprise,* Enugu, Nigeria, **2005.**

[16] URL: Http:// www.Mathworld.wolfram.com

[16] William J. Stevenson, *Operations Management,* McGraw Hill, New York. ISBN-10: 0077284097, ISBN-13: 978-0077284091, 10th Edition, **2008.**

Appendices

APPENDIX I

Sample of random numbers used to test for uniformity and independence.

(Inactive C:\TPW\HYPO.EXE)

RANDOM NUMBERS BETWEEN 0 AND 1

0.77377377	0.60860861
0.19519520	0.31831832
0.92592593	0.74674675
0.91391391	0.31831832
0.80980380	0.73373373
0.61961962	0.94294294
0.94594595	0.58358358
0.15615616	0.65765766
0.96996997	0.17117111
0.52352352	0.56056056
0.25325325	
0.71671672	
0.43043043	
0.71471471	
0.72672673	
0.97997998	
0.76276276	
0.14414414	
0.14314314	
0.69069069	