Contemporary Aspects of Monte Carlo Methods and Simulation in Physics

J A. OLOWOFELA, J . A .ADEGOKE, 0. P. ADEWUMI and I. C. KAMIYOLE
Department of Physics, University of Ibadan, Ibadan, Nigeria.
Department of Physics, Oyo-State College of Education, Oyo, Nigeria

Abstract

This article introduces Monte Carlo methods which are different from conventional numerical methods and show how some of the methods can be applied in Physics to simulate or solve physical problems, through computer programs (written in FORTRAN in this case), by using a few examples touching modern and classical Physics showing its usefulness. A functional approach to probability and statistics is described for the purpose of this work instead of complete treatment. The importance of good sequences of random numbers with large periods is demonstrated and the methods, in some instances are compared with a conventional type and differences pointed out.

History and Introduction

The expression "Monte Carlo" is actually very general and called so after a popular city of Monte Carlo in Europe, famous for its cosmopolitan character and gambling casinos, as a result of the stochastic nature of gambling and of the methods. They are based on the use of random numbers and probability statistics to investigate problems. One can find Monte Carlo methods used in everything from economics to nuclear physics. Monte Carlo methods have been used for centuries, but only in the past several decades has the technique gained the status of a full-fledged numerical method capable of addressing the most complex applications. The applications vary from field to field, and there are dozens of subsets of Monte Carlo, even within physics. To call something a "Monte Carlo" experiment, all one needs to do is use random numbers to examine some problems.

The use of Monte Carlo methods to model physical problems allows us to examine more complex systems than we otherwise can. Solving equations which describe the interactions for hundreds or thousands of atom is im possible but, with Monte Carlo methods, a large system can be sampled in a number of random configurations, and that data can be used to describe the system as a whole. Statistical simulation methods may be contrasted to conventional numerical discretisation methods , which typically are applied to ordinary or partial differential equations that describe some underlying physical or mathematical system. In many applications of Monte Carlo, the physical process is simulated directly, with the only requirement that the physical (or mathematical) system be described by probability density functions (pdf's) and no need to even write down the differential equations that describe the behaviour of the system. Many applications are then performed (multiple "trials") and the desired result is taken as an average over the number of observations. In many practical applications, one can predict the statistical error (variance) in this average result, and hence an estimate of the number of Monte Carlo trials that are needed to achieve a given error.

Monte Carlo methods necessitate a fast and effective way to generate random numbers uniformly distributed on the interval [The outcomes of these random samplings must be accumulated in an appropriate manner to produce the desired result, but the essential characteristics of Monte Carlo is the use of random sampling techniques (and perhaps other algebra to manipulate the outcomes) to arrive at a solution of the physical problem.

It is natural to think that Monte Carlo methods are used to simulate random processes, as these can be described by pdf's. However, many Monte Carlo applications have no apparent stochastic content, such as the evaluation of definite integral. In spite of this, one can pose the desired solution in terms of pdf's, and while this transformation may seem artificial, this step allows the system to be treated as a stochastic process for the purpose of simulation and hence Monte Carlo methods can be applied to simulate the system. One can, therefore, take a broad view of the definition of Monte Carlo methods by including rubric all methods that involve statistical simulation of some underlying system, whether or not the system represents a real physical process.

The major components of a Monte Carlo method, comprising the foundation of most Monte Carlo applications, are: (i) Probability distribution functions (pdf's), (ii) Random number generator, (iii) Sampling rule - rule for sampling from the specified pdf's must be given, (iv) Scoring- the out comes must be accumulated into overall tallies or scores for the quantities of interest (v) Error estimation - an estimate of the statistical error (variance) as a function of the number of trials and other quantities must be determined and (vi) Variance reduction techniques- this can lead to reduction in the computational time for Monte Carlo simulation as well.

What is intended to be shown in this article will be illustration of the application of some Monte Carlo Methods and their major components in the following areas (1) Simple Monte Carlo evaluation of an integral (2) Multidimensional Monte Carlo integration (3) Monte Carlo Error Analysis (4) The

Approach to Equilibrium - Monte Carlo Method, and (5)Radioactive Decay.

Random number generation and some Monte Carlo Methods.

Since Monte Carlo Methods are procedures that make use of random numbers, generation of random numbers of immense importance to the methods.

A sequence of random numbers is a set of numbers that have nothing to do with the other numbers in the sequence. The idea of a single random number is incorrect. In a uniform distribution of random numbers in the range [To be precise, the algorithm generate integers between 0 and, say M, and divide by M to return a real value within the range. An example called "middle square", describes generating a sequence of ten digit integers by starting with one, squaring it and then taking the middle ten digits from the answer as the next number in the sequence. The sequence is, however, not random, as each number is completely determined from the previous.

The power residue and linear congruent methods are, respectively, de scribed by equations:

$I = aI + c \text{ mod} M$ (1)

$I = (aIr + c) \text{mod} M$ (2)

The starting value (seed) of I is I where a, c and M are constants specifically chosen such that a and c and greater than or equal to 0 and M is greater than I a and c. A poor choice of the constants can lead to very poor sequences e.g. ones with short periods. The choice c=0, in the case of linear congruential method, obviously leads to a somewhat faster algorithm, and can also result in long sequences. This is called "Multiplicative congruential". M should be as large as possible since the period can never be longer than M. One's choice of M should be one near the largest integer that can be represented by the computer. A sample algorithm for the power residue method is given below:

STEP 1:specify the three constants to be used a,c and m and the seed. STEP 2:substitute them into the equation to obtain a value, say x.

STEP 3:if this is greater than m, set the value back to that of the seed.

STEP 4:whatever the case, divide the value x by m to get next number in sequence.

STEP 5: substitute into the seed's position the new x and return to step 2.

For any program making use of random number sequences it must be ensured that the amount of trials used does not exceed the periods of the sequences for reliable results to be obtained, other factors in order.

Two methods are demonstrated for the purpose of illustration by integrating

the function: 4* Vf— x within limits 0 and 1 with respect to x. The function is from the equation of the first quadrant of a circle of unit radius. The integer 4 in the expression ensures the area of the circle is what is obtained when the integration is carried out since there are four quadrants in any circle. It should be noted that the evaluation happens to be equivalent to the value of pi.

The first method. This method, can be viewed graphically such that the area of a particular function, which is to be determined, is represented within a larger, regular area or shape whose area is known and an algorithm developed so that ratio of number of randomly picked coordinates which fall within represented area of the function to total number of coordinates generated (which should always fall in those for the larger area) gives the area of the function when multiplied by area of the larger shape.

The evaluation of the function of the first quadrant (ABCD) (Figure 1) of a circle of unit radius within a square, using this method in the computer program algorithm, reveals that increase in trial numbers increases the ac curacy of the method. The table 1 of results for a particular sequence of random numbers of over 8 million counts as period is given below.

Value of pi using Hit-Miss Method.

Actual Value of pi=3.
14159265358979323846264338327 95

The logarithm of the errors in the value, when plotted against the total trials made, gives a graph (figure 2) showing that the greater the total trial number, the better the accuracy, that is, for large trial total, say N, error approaches zero as inverse of square root of N. This is revealed from the values "DEP", which is the ratio of the error in the value of pi to the trial total N, calculated for two instances as seen in the result. It establishes the relationship between error and trial total.

The second and final method for the purpose of this article is the "Sample Mean" method (Figure 2). In definite integration, with limits a and b manipulation is made such that a probability distribution function f(x) is defined to be 1/(b-a) for a < x < and zero for any other values and introduced into the integral (by dividing the original function) with the original integral viewed as the expectation value of the function in it so that a final expression of the for:

$(b-a)/N*4 \_x$ (3)

is obtained, where b=1 and a=0 for this particular example. USING SAMPLE MEAN METHOD (Table 2).

It can be seen that the values fluctuate for small trial totals but steadily approach the true value as trials increase. With trials of 1,000,000, accuracy to 2 decimal places can be said to be achieved and higher accuracies obtained with increase in trial total (N). By dividing variance by total trial number ("DEP") and the fact that there is correspondence between error and variance, one can observe, as shown by the last four lines of data in the in the result, that the approximate functional dependence of error on N is such that error approaches zero as inverse of square root N, for large N. When a trial total is squared, square root of its error is obtained. "ERROR 2" show ap proximate deviation from true pi obtained by

dividing standard deviation by square root of trial total.

When the lines of algorithm or actual code lines used in the two methods are compared the sample mean method is seen to take a shorter time to exe cute and this, in addition to the fact that it gives better result for the same number of trials, makes it a better choice. Typically, the hit-miss method takes more than twice the computational time required by the sample mean method for a considerable trial total.

Simpson's rule, a conventional numerical method, requires lots partitions, or iterations, to be able to obtain a very accurate value of pi. A particular result is shown below

VALUE OF PT USING SIMPSON'S RULE:
PARTITIONS MADE = 119
STMPSONS GIVES P1= 3.138611

However, it gives a good result for a small partition total of 119, comparing with values for the trial of 1000 obtained in the previous methods. Judging from the code needed to have used Simpson's rule, it can be considered to be inefficient with computer time compared with the previous Monte Carlo methods.

The choice of using one method instead of the other then depends on the objectives of the programmer. It would however be a bad choice to use conventional methods for multidimensional integration as the error introduced
would increase and considerable computer time will be wasted.

3 Error Analysis

This involves a way of obtaining the error present from any evaluation so that the degree of accuracy of a particular method used can be verified. This will be demonstrated by integrating the function mentioned earlier for evaluating pi, taking trial to be 10,000 trials, and obtaining the standard deviation of the measurements. The second is by dividing the 10,000 trials into sub-sets and the standard deviation also obtained of the measurements. The second is by dividing the 10,000 trials into sub-sets and the standard deviation also obtained. In the first case, the question "How can one tell if only 10 measurements will give the desired accuracy?" still remains, and, appropriately, can one tell whether 10 sub-divisions are appropriate in the second case? These suggested trying various options and analysing the results. It should be stated that, as observed from previous work, standard deviation of a single measurement is incorrect as error in value of pi obtained at any time, hence the need for the above procedures.

It is worth remembering that the variance, square of standard deviation, of m measurements is given by:

(4)

where

(5)

(6)

Results to demonstrate the first procedure reveal that standard deviation of each set of measurements does not give the correct error (Table 3), but values obtained in Table 4 are consistent with the error obtained in previ ous evaluations where actual pi value minus calculated ones gave the error involved. Also Table 4 shows that any set of measurements can be appropriate by saving computer time. The interpretation of the standard deviation for, say 10 measurements, is that F has 68.36% chance of being within 0.006836 of the true mean, where F is the estimate.

It is observed, from Table 4, that sets of measurement above 20 give larger standard deviation, but all have more or less same "2ND MEASURE OF ERROR".

"2ND MEASURE" gives values of errors determined for each in another way. It is by dividing standard deviation for a single measurement in a set by the square root of trial total (10,000 in this case), i.e o It should be noted that "ST.DEVIATION" of Table 4 is calculated from the mean of the number of sets used, each set having 10 measurements.

With the second procedure similar steps are taken but trial total (10,000) is divided into equal partitions. The standard deviation of the sub-divisions is given by square root of the expression below:

(7)

As expected, from the third table, standard deviation of the functions of a single division does not give accurate error involved. Table 4 reveals that the standard deviation of each set of sub-divisions is also inconsistent with previous errors, but the expressions gives consistency and is the error involved recorded as "ERROR".

It is observed that 10 sub-sets might be the preferable choice of sub-divisions, as 20 has higher probable error and picking most above 20 only take computer time without improving on the error.

To drive home the point of using a good sequence of random numbers, a bad sequence of period of period two is generated and the results analysed taking the following constants: 5, 0 and 32 as a, c and m. The power residue method is used for this purpose. First, to calculate pi for varying trial numbers, it is obvious that the values obtained for pi are off the mark, and even got worse for large total trials, as demonstrated by sample Mean method and the integral used previously.

The two procedures mentioned earlier are also affected as values turned deterministic. For the first procedure (10,000 trials each), it can be seen that the values of pi repeat themselves with the same period, two, as the sequence used and Table 6 couldn't be generated (unlike before) as variance of the measurements for each group is negative which wouldn't give real standard deviations.

Similar problems (error in value of pi and repetition are observed when the second procedure is used and its Table 10 shows "calculated error" for only two sets of subdivisions, since the variance of the remaining yield negative values.

Application to Some Physical Examples

The Approach to Equilibrium

An ideal gas of identical, distinguishable particles in a box for which inter action might be ignored and the box is isolated is considered. The initial state of the system is such that all particles are in the left half of the box and distribution occurs when a small slide, covering a hole in the middle, the point of division of the box in two, is removed. One assumption made is that one particle passes though the hole in a unit time. This section answers questions like, "at time t, what is the average number of particles left in the left half of the box?". "What is the probable equilibration time?". This illustrations demonstrates for total particles of 10, 20 and 40, three different systems. The equilibrium state is achieved by a system when, after some time, its macroscopic state does not change in time except for small random fluctuations.

The algorithm of the program is such that it ensures only the first time when the particles are equally spread out is recorded as equilibration time since only small fluctuations follow later. A sample algorithm is given below:

STEP 1: Specify the constant initial amount in the left half of the box.

STEP 2: Allow appreciable time for equilibration, say, ten times amount.

STEP 3: Initialize variable to hold equilibration time

STEP 4: Find the ratio of particles in the left half to total and compare with random number

STEP 5: If ratio is less than or equal, then a particle moves to the right, else, it moves left.

STEP 6: Repeat steps 4 and 5, each time checking when, first, half of the total remains in the left until time allowed elapse

STEP 7: Another system may be worked for, say, doubling the former, and step 2 is returned to:

The results below are for a particular sequence of random numbers.

PARTICLE-TOTAL = 10
PARTICLE LEFT IN LEFT PART 4
APPR. TIME(S) TO REACH EQUILIBRIUM = 9
PARTICLE-TOTAL = 20
PARTICLE LEFT IN LEFT PART = 10
APPR. TIME(S) TO REACH EQUILIBRIUM = 32
PARTICLE-TOTAL = 40
PARTICLE LEFT IN LEFT PART = 18
APPR. TIME(S) TO REACH EQUILIBRIUM = 50

It is observed that the system approached equilibrium (Figure 3) and for more initial, total particles it took longer time to approach equilibrium (Figures 4 and 5), as expected. Since the tendency in nature is towards disorder (equilibration, in this case)

The next graph (Figure 6) was generated by a program written to compare the approach to equilibrium for five different trials and their averaged value using constant total particles of 10. Again their initial position is the left half of the box and other assumptions are as in the earlier example. Hence the method of achieving results is similar to the former. The relative magnitude of the fluctuations can be determined for each unit time after equilibrium is reached, where " " represents a time average, the standard deviation a is the square root of (<n > — < >2) and n is the average number of particles in the left half of the box.

The graph of the averaged value (Figure 6) is seen to lie mid-way as expected.

Radioactive Decay

Radioactive decay is truly a random process and it provided the first evidence that the laws that govern the sub-atomic world are statistical. With the choice of a system of initial, parent particle number of 100, decay constant 0.01 and duration of 300 seconds, radioactive decay was simulated using the sequence of random numbers generated by the computer. Two simulations

were performed. The first compares the random numbers with the constant probability of decay,) of a particle (where \) is the decay constant and L is in 1 second in this case) with the assumption that each particle has constant probability of decay per unit time, and this probability must be much lesser than 1. While the second uses poisson distribution, with random numbers compared with varying probabilities of decay, NA At, where each particle has constant probability and N is the parent particle number left after time t. The Poisson distribution procedure involved dividing, by 3,000 in this case, the duration into smaller periods ensuring the probability for is a lot lesser than 1, as a rule. The probability of observing, say, n decays is P = P — P)"' rn for m divisions. Where p is the probability of a single decay per division and C represents Combination.

HITS	TRIALS	CIRCLE AREA/PI	ERROR
771	1000	3.08400011	0.057559263
1555	2000	3.10999990	0.03159285
2378	3000	3.17066669	−0.02907395
3144	4000	3.14400005	0.00240731
3903	5000	3.12240005	0.01919270
4678	6000	3.11866665	0.02292609
5476	7000	3.12914276	0.01244998
6300	8000	3.15000010	−0.00840735
7060	9000	3.13777781	0.00381494
7884	10000	3.15359998	0.01200724

N−1000 AREA−3.08400011 DEP−0.0000575926
N−1000,000 AREA=3.14289999 DEP=0.0000000013

Table 2

SET	TRIALS	AREA	STANDARD DEVIATION	ERROR2
1	1000	3.1432	0.903691	0.028577
2	2000	3.1410	0.892880	0.019965
3	3000	3.1170	0.905188	0.016526
4	4000	3.1188	0.906927	0.014340
5	5000	3.1286	0.902419	0.012762
6	6000	3.1448	0.890570	0.011497
7	7000	3.1446	0.891380	0.010654
8	8000	3.1440	0.891746	0.009970
9	9000	3.1430	0.895645	0.009441
10	1000	3.1416	0.897257	0.008973

N=1000 PI=3.1432 VAR.=0.816658 DEP=0.00081666 ERR 2 = 0.028577
N=1000,000 PI=3.1467 VAR.=0.783971 DEP=0.00000078 ERR 2 =0.00085

Table 3
Error Analysis for 10 Measurements
of 10,000 Trials each

SET	PI	STD.DEVIATION
1	3.1405	0.896342
2	3.1446	0.894922
3	3.1337	0.896822
4	3.1405	0.890015
5	3.1394	0.901376
6	3.1302	0.906212
7	3.1461	0.887876
8	3.1294	0.892418
9	3.1510	0.887860
10	3.1473	0.881664

Table 4
Test of Sets of Measurements

NO OF SETS	PI	STANDARD DEVIATION	2ND MEASURE
10	3.140263	0.006836	0.008817
20	3.144753	0.007308	0.008963
30	3.143187	0.010242	0.008860
40	3.144230	0.009863	0.008783
50	3.140771	0.008843	0.009051
60	3.141572	0.010472	0.008948
70	3.141069	0.009418	0.008895
80	3.140614	0.009618	0.009055
90	3.140763	0.009213	0.008845
100	3.141583	0.009213	0.009026

Table 5
Error Analysis for 10,000 Trials
divided into 10 Sub-Divisions

SET	PI	STD.DEVIATION
1	3.12987	0.885163
2	3.12549	0.896554
3	3.12378	0.900514
4	3.10540	0.928437
5	3.09956	0.919046
6	3.14333	0.924043
7	3.17296	0.853064
8	3.16376	0.859379
9	3.17371	0.902643
10	3.16699	0.887293

Table 6
Test of Sub-Divided Measurements
divided into 10 Sub-Divisions

DIVISIONS	TRIALS	PI	ST.DEV.	ERROR
10	1000	3.14049	0.026349	0.008332
20	500	3.14463	0.052807	0.011808
30	333	3.13359	0.040853	0.007459
40	250	3.14056	0.052951	0.008372
50	200	3.13914	0.076028	0.010752
60	166	3.13139	0.066032	0.008525
70	142	3.14371	0.078806	0.009419
80	125	3.13105	0.085247	0.009531
90	111	3.15135	0.084398	0.008896
100	100	3.14625	0.081897	0.008190

Table 7
Using Sample Mean Method with Sequence of Period 2.

SET	*TRIALS*	*PI*	*ST.DEV.*	*ERROR*
1	1000	3.4824	0.696599	0.022028
2	2000	3.4817	0.696884	0.015583
3	3000	3.4820	0.696614	0.012718
4	4000	3.4821	0.696406	0.011011
5	5000	3.4819	0.696692	0.009850
6	6000	3.4820	0.696645	0.008994
7	7000	3.4820	0.696754	0.008328
8	8000	3.4819	0.696908	0.007792
9	9000	3.4819	0.696964	0.007347
10	10000	3.4819	0.697391	0.006974

= 10,000 Pi=3.4825 Std. Dev. = 0.702892 Err.= 0.002223
= 100000 Pi= 3.4941 Std.Dev. = 0.617401 Err.= 0.000617

Table 8
Error Analysis for 10 measurements of 10,000 Trials each Using Sequence of Period 2

SET	*PI*	*STDDEVIATION*
1	3.481856	0.69731
2	3.481710	0.697445
3	3.481861	0.697395
4	3.481856	0.697391
5	3.481710	0.697445
6	3.481861	0.697395
7	3.481856	0.697391
8	3.481710	0.697445
9	3.481861	0.697395
10	3.481856	0.697391

Table 9
Error Analysis for 10,000 Trials Divided into
10 Measurements for Sequence of Period 2

SET	*PI*	*STD.DEVIATION*
1	3.48240	0.696599
2	3.48245	0.696633
3	3.48240	0.696599
4	3.48240	0.696599
5	3.48095	0.697136
6	3.48245	0.696633
7	3.48240	0.696599
8	3.48095	0.697136
9	3.48245	0.696633
10	3.48240	0.696599

Table 10
Table 10: Test of Sub-Divided Measurements

DIVISIONS	*TRIALS*	*STD.DEVIATION*	*ERROR*
1	1000	0.001381	0.000437
2	500	0.002184	0.000488

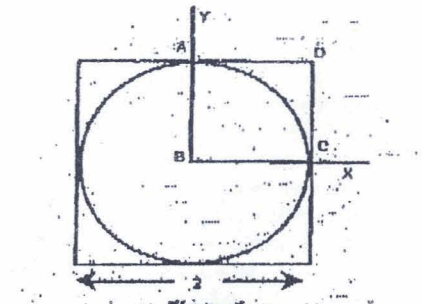
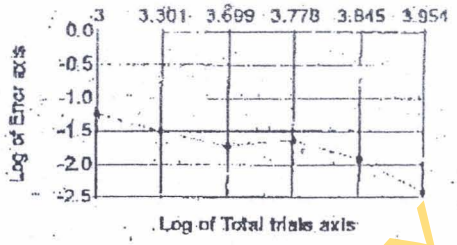


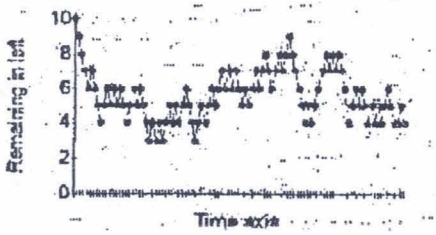
Figure 1
GEOMETRY FOR THE HIT-MISS METHOD TO FIND THE VALUE OF PI OR THE AREA OF A CIRCLE OF UNIT RADIUS.

FIGURE 2

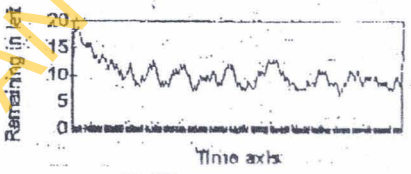


Only six points appear as log of negative values cannot be taken

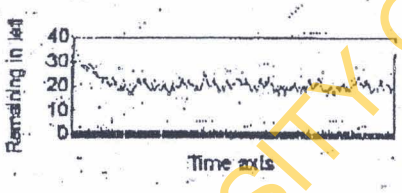
Approach to equilibrium, 10 particles
FIGURE 3



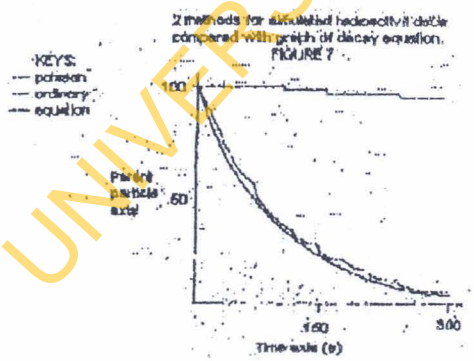
Approach to equilibrium, 20 particles
FIGURE 4



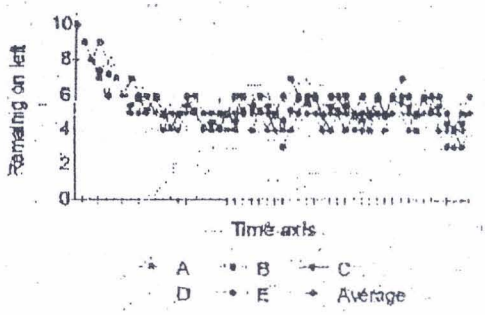
Approach to equilibrium, 40 particles
FIGURE 5



2 methods for simulated radioactivity decay compared with graph of decay equation.
FIGURE 7



Approach to equilibrium, 5 trials
FIGURE 8



Graphs of both were plotted and a third graph showed the pattern obtained from the equation of radioactive decay N(t) = Noe_)\t.

A little trick was used to draw the decay graph (Figure 7) of the poisson distribution procedure on the same scale as the others, an averaged particle for every ten steps was taken and recorded as number of

parent particles left per unit time. It worked because of the large value 3,000 used in dividing.

The result (Figure 7) shows the mean and variance for the exponential distribution of which the decay equation is an example, the variance being the square of the mean which is itself the inverse of the decay constant. The same sequence was used to generate the simulated decays but the one through "unaided" application deviated quite differently. Hence the need to generate random numbers according to particular distribution in some cases. The Poisson distribution, however, faired much better with its graph and report on remaining prent particles; as confirmed by calculation through the decay formula.

Conclusion

Working with Monte Carlo methods can be easy, representative of the de sired process or operation and fun but requires great care in some instances. This is because not the same sequence of random numbers can be applied in all physical processes and the same degree of accuracy expected in all cases. The way most, if not all, computers generate random numbers is such that at the end of a certain program which uses random numbers, when a new program which uses random numbers as well, is started it makes use of the same sequence used by the earlier program picking each number in the same order.

Special sequences can be developed (sequences according to specific distributions) for different processes or operations so that better results are achieved. Special purpose algorithms (found in numerical libraries) are used in this situation. General purpose method according to an arbitrary distribution function can also be used instead, an example being "rejection technique", i.e. Hit-Miss method, which was described earlier. However this method is not suitable for distribution with one or more large peaks. Another method which is suitable for relatively simple distribution functions is the "inversion technique". The inversion techniques has the following steps: (1) Normalisation of the distribution function to get a pdf, (2) integration of the pdf analytically from the minimum x to an arbitrary x, representing the prob ability of choosing a value less than x and (3) Equating this to a uniform number and solving for x. The resulting x will be distributed according to the pdf. This method is considered efficient.

References

Harvey Gould and Jan Tobochink (1988): An Introduction to Computer

Simulations: Applications to Physical Systems (Part 1 2) Addison-Wesley Publishing Company

Sen, S.K. (1967): Simulation of the Rayleigh-Pearson random walk, J. nd. Inst. Sci,49,19-27

Malvin H Kalos and Paula A. Whitlock (1986): Monte Carlo Methods, Vol. 1: Basics, John Wiley Sons.

Reuven Y.Rubinstein (1981).Simulation and the Monte Carlo Method John Willey Sons, Computational Physics

John R. Taylor (1982),An Introduction to Error Analysis, University Science Books, Oxford University Press.