JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 64B, No. 4, October–December 1960

Accuracy of Monte Carlo Methods in Computing Finite Markov Chains

N. W. Bazley and P. J. Davis

(May 20, 1960)

Experiments are made with the Markov chain presented by the children's game of Chutes and Ladders. Statistics, such as the average length of play, are computed on the IBM 704 from 2^{14} simulated plays of the game. These Monte Carlo results are then compared with the "exact solution" obtained by powering the matrix of transition probabilities. Convergence is shown to obey the familiar "$N^{-1/2}$" law.

1. Introduction

Since Monte Carlo methods are generally applied to problems that are too difficult for other types of analysis there is seldom opportunity to compare their results with "exact solutions." In this paper experiments are described with the Markov chain presented by the children's game of Chutes and Ladders.[1] In particular, the calculation of the average length $\overline{L}$ of a play of this game is considered. An exact value for $\overline{L}$ is found according to Markov chain theory. The Monte Carlo calculation for $\overline{L}$ involves 2^{14} simulated plays of the game. Furthermore, the probabilities of being in a given state by a fixed number of throws are computed theoretically and experimentally.

It is shown that well-known theorems of statistics imply that the Monte Carlo results should converge to the true solutions by "$N^{-½}$" laws. Numerical values are shown to follow these laws. All work was carried out on the IBM 704 computer.

2. Description of the Game

First the game of Chutes and Ladders is described in detail. The game is considered to be played in its solitaire form. The game involves a playing piece, one die, and a playing board which has squares numbered from 1 to 100. Certain pairs of these 100 squares are connected by "chutes" and others are connected by "ladders." The purpose of the chutes is to delay the game by forcing the player to return to the square at the bottom of a chute if he should land on the square at its top. On the other hand, the ladders speed up the game by permitting a player to advance directly from the square at the bottom of a ladder to the square at its top, and thus bypass the squares in between. There are 10 chutes connecting the pairs of numbered squares (16,6), (47,26), (49,11), (56,53), (62,19), (64,60), (87,24), (93,73), (95,75), and (98,78). However, nine ladders connect the squares (1,38), (4,14), (9,31), (21,42), (28,84), (36,44), (51,67), (71,91), and (80,100).

Since the playing piece can never remain at either the top of a chute or the bottom of a ladder, there are only 81 squares at which it can stop.

The playing piece is off the board at the start of the game and the player tosses the die so as to determine a number from 1 to 6. He then moves his playing piece through the same number of squares as the number on the die. Should his playing piece stop at the bottom of a ladder he must immediately move to the top of the ladder. He then repeats the throw of the die and moves accordingly. If at any time he lands at the top of a chute he must proceed directly to its bottom before the next die toss. The game is concluded when the playing piece reaches square 100. The board actually has two winning positions, since there is a ladder from square 80 to square 100. One requirement is that square 100 be reached by an exact throw of the die.

3. Theoretical Solution

The game can be interpreted as a finite Markov chain with 82 states. These states are indexed by $i=0, 1, \ldots, 81$, where the state $i=0$ corresponds to the starting position and the state $i=81$ corresponds to the playing piece being at square 100. All squares except those at the top of a chute or the bottom of a ladder have a corresponding state in the chain.

Let $X_i(n)$, $i=0, 1, \ldots, 81$ represent the probability of being in state i after n throws of the die. Then $X_0(0)=1$ and $X_0(n)=0$ for $n>0$. Also, $\sum_{i=0}^{81} X_i(n)=1$ for every $n\geq 0$. Let p_{ij} be the conditional probability of moving from square i to square j on one die toss whenever square i is occupied. We see immediately that

$$\begin{bmatrix}\text{Prob of state}\\ j \text{ on } n \text{ tosses}\end{bmatrix}=$$

$$\sum_{i=0}^{81}\begin{bmatrix}\text{Prob of transition}\\ \text{from state } i \text{ to}\\ \text{state } j \text{ on one toss}\end{bmatrix}\cdot\begin{bmatrix}\text{Prob of state } i\\ \text{on } n-1 \text{ tosses}\end{bmatrix}\cdot \quad (3.1)$$

[1] Copyright Milton-Bradley Co.

In other words, the elements $X_j(n)$ satisfy the relations

$$X_j(n)=\sum_{i=0}^{81} p_{ij}X_i(n-1), \qquad \begin{array}{l} 0\leq j\leq 81 \\ n=1,2,3, \ldots \end{array} \tag{3.2}$$

where $X_i(0)=\begin{cases} 0, & 1\leq i\leq 81 \\ 1, & i=0 \end{cases}$. If we let $X(n)$ be a row vector whose components are $X_j(n)$ we can write (3.2) in the form

$$X(n)=X(n-1)P=X(0)P^n \tag{3.3}$$

where $X(0)$ is the vector whose components are $X_0(0), X_1(0), \ldots, X_{81}(0)$ and P^n is the nth power of

$$P=\begin{bmatrix} p_{00}, & p_{01}, & \ldots, & p_{0,81} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ p_{81,0}, & p_{81,1}, & \ldots, & p_{81,81} \end{bmatrix}.$$

The matrix elements p_{ij} will be called the transition probabilities; certain of them are given in table 1.

We now consider the average length, $\bar{L}$, of the game. This may be computed from the expression

$$\bar{L}=\sum_{n=1}^{\infty} n[X_{81}(n)-X_{81}(n-1)], \tag{3.4}$$

where the term $X_{81}(n)-X_{81}(n-1)$ represents the probability of reaching the final square on die toss number n.

An alternative expression which can be used for computing $\bar{L}$ is given as follows. First introduce the truncated matrix P_T formed by crossing out both the last row and the last column of P. Then form the matrix

$$T=(I-P_T)^{-1} \tag{3.5}$$

TABLE 1. *Partial listing of matrix elements p_{ij}*

i \ j	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0		1/6	1/6	1/6	1/6							1/6			
1			1/6	1/6	1/6	1/6	1/6					1/6			
2				1/6	1/6	1/6	1/6					1/6			
3					1/6	1/6	1/6	1/6	1/6						
4						1/6	1/6	1/6	1/6	1/6					
5							1/6	1/6	1/6	1/6	1/6				
6								1/6	1/6	1/6	1/6	1/6			
7				1/6					1/6	1/6	1/6	1/6	1/6		
8				1/6						1/6	1/6	1/6	1/6	1/6	
9				1/6							1/6	1/6	1/6	1/6	1/6
10				1/6								1/6	1/6	1/6	1/6
11				1/6									1/6	1/6	1/6
12				1/6										1/6	1/6
13															1/6
14															
. . .															

i \ j	69	70	71	72	73	74	75	76	77	78	79	80	81
66	1/6	1/6	1/6										
67	1/6	1/6	1/6	1/6									
68	1/6	1/6	1/6	1/6	1/6								
69		1/6	1/6	1/6	1/6	1/6							
70			1/6	1/6	1/6	1/6	1/6						
71				1/6	1/6	1/6	1/6	1/6					
72					1/6	1/6	1/6	1/6	1/6				
73						1/6	1/6	1/6	1/6				
74							1/6	1/6	1/6	1/6			
75								1/6	1/6	1/6	1/6		
76									1/6	1/6	1/6		
77										1/6	1/6	1/6	1/6
78										1/3	1/6	1/6	1/6
79											1/2	1/6	1/6
80												5/6	1/6
81													1

with elements T_{ij}. It is shown on p. 329 of Kemeny, Snell, and Thompson [2] that $\bar{L}$ is given by

$$\bar{L}=\sum_{j=1}^{80} T_{0j}, \tag{3.6}$$

which is simply the sum of the first row of the matrix $(I-P_T)^{-1}$.

Equation (3.3) shows that a direct numerical solution by successive matrix vector multiplications can be used to calculate the vectors $X(n)$. This solution may not be feasible for full matrices or for chains with a large number of states, since powering the matrix may be subject to space and time limitations and roundoff error. However, for Chutes and Ladders only 476 of the 6,724 transition probabilities do not vanish. The $X(n)$ were easily calculated on the IBM computer by taking the sums in (3.2) over the nonvanishing elements. The time required for each matrix vector multiplication was about 2 sec. To illustrate the results we have presented the vectors $X(7)$ and $X(31)$ in table 2. The vector $X(7)$ is of interest since the computations showed that $X_{81}(n)=0$ for $n<7$, so that the game cannot be completed in fewer than seven throws; the vector $X(31)$ is of interest since $X_{81}(31)$ has also been computed by Monte Carlo.

TABLE 2. *The vectors $X(7)$ and $X(31)$*

i	$X_i(7)$	$X_i(31)$	i	$X_i(7)$	$X_i(31)$
0	0	0	41	$.15936\times10^{-1}$	$.62209\times10^{-2}$
1	0	0	42	$.88735\times10^{-2}$	$.37065\times10^{-2}$
2	0	0	43	$.16175\times10^{-1}$	$.74137\times10^{-2}$
3	0	0	44	$.95950\times10^{-2}$	$.42890\times10^{-2}$
4	$.19740\times10^{-1}$	$.32552\times10^{-2}$	45	$.79733\times10^{-2}$	$.37599\times10^{-2}$
5	$.47725\times10^{-2}$	$.56622\times10^{-3}$	46	$.71266\times10^{-2}$	$.33331\times10^{-2}$
6	$.60942\times10^{-2}$	$.66476\times10^{-3}$	47	$.81697\times10^{-2}$	$.39130\times10^{-2}$
7	$.76982\times10^{-2}$	$.78046\times10^{-3}$	48	$.78589\times10^{-2}$	$.39497\times10^{-2}$
8	$.28607\times10^{-1}$	$.81893\times10^{-2}$	49	$.92057\times10^{-2}$	$.73498\times10^{-2}$
9	$.14518\times10^{-1}$	$.23394\times10^{-2}$	50	$.50654\times10^{-2}$	$.38812\times10^{-2}$
10	$.13064\times10^{-1}$	$.21801\times10^{-2}$	51	$.40509\times10^{-2}$	$.39027\times10^{-2}$
11	$.14489\times10^{-1}$	$.24607\times10^{-2}$	52	$.20898\times10^{-2}$	$.33210\times10^{-2}$
12	$.16036\times10^{-1}$	$.27729\times10^{-2}$	53	$.12610\times10^{-2}$	$.32115\times10^{-2}$
13	$.17811\times10^{-1}$	$.31194\times10^{-2}$	54	$.12867\times10^{-1}$	$.74863\times10^{-2}$
14	$.17300\times10^{-1}$	$.22388\times10^{-2}$	55	$.21219\times10^{-2}$	$.31170\times10^{-2}$
15	$.22630\times10^{-1}$	$.61242\times10^{-2}$	56	$.24970\times10^{-2}$	$.36590\times10^{-2}$
16	$.21066\times10^{-1}$	$.29082\times10^{-2}$	57	$.27328\times10^{-2}$	$.36162\times10^{-2}$
17	$.19726\times10^{-1}$	$.25041\times10^{-2}$	58	$.31579\times10^{-2}$	$.36672\times10^{-2}$
18	$.24366\times10^{-1}$	$.28399\times10^{-2}$	59	$.71766\times10^{-2}$	$.11658\times10^{-1}$
19	$.30696\times10^{-1}$	$.12958\times10^{-1}$	60	$.26578\times10^{-2}$	$.44668\times10^{-2}$
20	$.26285\times10^{-1}$	$.47659\times10^{-2}$	61	$.55870\times10^{-2}$	$.12502\times10^{-1}$
21	$.50694\times10^{-1}$	$.12747\times10^{-1}$	62	$.25184\times10^{-2}$	$.62322\times10^{-2}$
22	$.27156\times10^{-1}$	$.62357\times10^{-2}$	63	$.21969\times10^{-2}$	$.66853\times10^{-2}$
23	$.26149\times10^{-1}$	$.68824\times10^{-2}$	64	$.30328\times10^{-2}$	$.14594\times10^{-1}$
24	$.24306\times10^{-1}$	$.75654\times10^{-2}$	65	$.18754\times10^{-2}$	$.97370\times10^{-2}$
25	$.29653\times10^{-1}$	$.74122\times10^{-2}$	66	$.82519\times10^{-3}$	$.86279\times10^{-2}$
26	$.22534\times10^{-1}$	$.70915\times10^{-2}$	67	$.42867\times10^{-3}$	$.79559\times10^{-2}$
27	$.18558\times10^{-1}$	$.61091\times10^{-2}$	68	$.26792\times10^{-3}$	$.82531\times10^{-2}$
28	$.17693\times10^{-1}$	$.60876\times10^{-2}$	69	$.30950\times10^{-1}$	$.15841\times10^{-1}$
29	$.20294\times10^{-1}$	$.71449\times10^{-2}$	70	$.42081\times10^{-2}$	$.87403\times10^{-2}$
30	$.16547\times10^{-1}$	$.58783\times10^{-2}$	71	$.47404\times10^{-2}$	$.85657\times10^{-2}$
31	$.14025\times10^{-1}$	$.56122\times10^{-2}$	72	$.53905\times10^{-2}$	$.85535\times10^{-2}$
32	$.12767\times10^{-1}$	$.53557\times10^{-2}$	73	$.61050\times10^{-2}$	$.86556\times10^{-2}$
33	$.11956\times10^{-1}$	$.52252\times10^{-2}$	74	$.68837\times10^{-2}$	$.87243\times10^{-2}$
34	$.11335\times10^{-1}$	$.50757\times10^{-2}$	75	$.67873\times10^{-2}$	$.11733\times10^{-1}$
35	$.33790\times10^{-1}$	$.77028\times10^{-2}$	76	$.38437\times10^{-2}$	$.80086\times10^{-2}$
36	$.17929\times10^{-1}$	$.60554\times10^{-2}$	77	$.35044\times10^{-2}$	$.79118\times10^{-2}$
37	$.38334\times10^{-1}$	$.13278\times10^{-1}$	78	$.24934\times10^{-2}$	$.96411\times10^{-2}$
38	$.22394\times10^{-1}$	$.74177\times10^{-2}$	79	$.16932\times10^{-2}$	$.13428\times10^{-1}$
39	$.23827\times10^{-1}$	$.77759\times10^{-2}$	80	$.41795\times10^{-3}$	$.34892\times10^{-1}$
40	$.23309\times10^{-1}$	$.73372\times10^{-2}$	81	$.15111\times10^{-2}$	$.48004$

[2] J. Kemeny, J. Snell, and G. Thompson, An introduction to finite mathematics, (Prentice-Hall, Inc., New York, N.Y., 1957).

The number $\bar{L}$ was computed from eq (3.5) and gave

$$\bar{L}=39.22 \tag{3.7}$$

as the average length of the game. The same result was obtained by summing the series in (3.6). The machine computing time was 12 min for the former calculation and 15 min for the latter.

4. Simulations

The random element in the game of Chutes and Ladders arises from the toss of a die at each move. This was simulated in our work by employing the pseudo-random number generator of Taussky and Todd [3]. That is, a sequence of pseudo-random numbers $t_1, t_2, \ldots$, where $0 \leq t_i \leq 1$, were generated by the rules

$$\left.\begin{aligned} t_{j+1}&=2^{-35}r_j \\ r_{j+1}&=5^{15}r_j \pmod{2^{35}}, \quad r_0=1. \end{aligned}\right\} \tag{4.1}$$

The unit interval was divided into six equal subintervals and the result of the jth die throw was identified with the subinterval (first, second, . . ., sixth) in which t_j fell. The game was then played N times on the calculator. The total number of throws necessary to play these N games was recorded as $\sum_{i=1}^{N} L_i$, where L_i denotes the length of each game. The average length of the game was then computed as $(1/N)\sum_{i=1}^{N} L_i$. Furthermore, the fraction of these games which finished in 31 or less throws was computed. For $N=2^{14}$ the computing time was about 45 min. The results of these calculations are given in table 3.

TABLE 3. *Results of the Monte Carlo calculations*

N	$\frac{1}{N}\sum_{i=1}^{N} L_i$	Fraction of games finished by 31 throws
32	37.469	.56250
64	36.969	.57813
128	36.117	.55469
256	39.480	.46484
512	38.701	.49414
1024	39.521	.47168
2048	39.032	.49072
4096	39.288	.47974
8192	39.469	.47693
16384	39.420	.47406
Exact answer	39.224	.48004

5. Convergence

We first show how $(1/N)\sum_{i=1}^{N} L_i$ should converge to $\bar{L}=39.22$. We introduce the distribution function

$$f(n)=X_{81}(n)-X_{81}(n-1). \tag{5.1}$$

[3] O. Taussky and J. Todd, Symposium on Monte Carlo methods (John Wiley & Sons, New York, N.Y., 1956).

The mean of $f(n)$ is simply the average length of the game $\overline{L}$ as given by eq (3.4). Furthermore, the standard deviation for the distribution function $f(n)$ is given by

$$\sigma=\sum_{n=1}^{\infty}(n-\overline{L})^2[X_{81}(n)-X_{81}(n-1)]. \tag{5.2}$$

The value of σ for Chutes and Ladders was calculated according to (5.2) with $\overline{L}=39.224$ and the known values of $X_{81}(n)-X_{81}(n-1)$. We found $\sigma=25.222$.

Let $L_1, L_2, \ldots, L_N$ be regarded as a random sample of size N drawn from the population with distribution $f(n)$ where the moment generating function of $f(n)$ exists. It is well known [4] that even though $f(n)$ is not normally distributed, the "sample mean" $1/N\sum_{i=1}^{N}L_i$ has a population distribution which approaches a normal distribution with mean $\overline{L}$ and standard deviation $\sigma/\sqrt{N}$ as N becomes infinite. Hence the probability is approximately .95 that the average length of play computed from the first N simulations lies in an interval centered at $\overline{L}$ with length $4\sigma/\sqrt{N}$. This bound involves σ, which has been computed in the case of Chutes and Ladders.

In figure 1 the theoretical values of $\overline{L}$ and σ are used to plot the interval $(\overline{L}-2\sigma/\sqrt{N}, \overline{L}+2\sigma/\sqrt{N})$ as a function of N. The values of $(1/N)\sum_{i=1}^{N}L_i$ are also plotted and it is seen that they lie within our confidence interval.

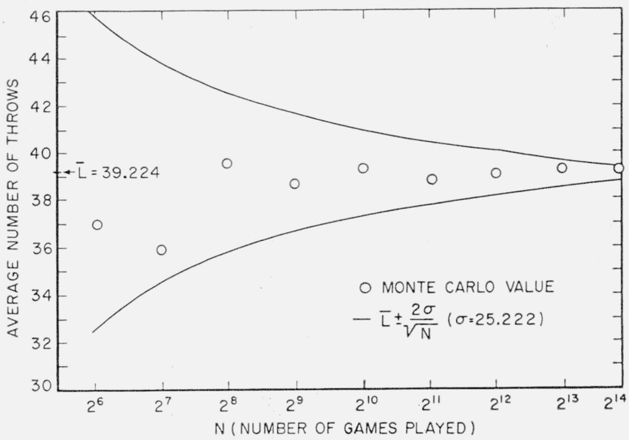

FIGURE 1. *Convergence to $\overline{L}$ by "$N^{-1/2}$ law."*

Now consider the probability of completing the game by a specified number of moves, say 31. If $p=X_{81}(31)$, then the first N simulations can be regarded as N Bernoulli trials for which the probability of success in a single trial is p. The probability of exactly ν successes is given by the binomial distribution

$$\frac{N!}{\nu!(N-\nu)!}p^{\nu}(1-p)^{N-\nu} \qquad (p=X_{81}(31), \quad 0\leq\nu\leq N). \tag{5.3}$$

For large N, however, the fraction ν/N of successes is approximately normally distributed with mean p and standard deviation $\sqrt{p(1-p)/N}$. Approximately 95 percent of the area under this normal curve lies in the interval

$$p-\frac{2\sqrt{p(1-p)}}{\sqrt{N}}<X<p+\frac{2\sqrt{p(1-p)}}{\sqrt{N}}, \tag{5.4}$$

and at least 95 percent lies in the larger interval

$$p-\frac{1}{\sqrt{N}}<X<p+\frac{1}{\sqrt{N}}. \tag{5.5}$$

Thus p can be estimated from ν/N, as is illustrated in figure 2.

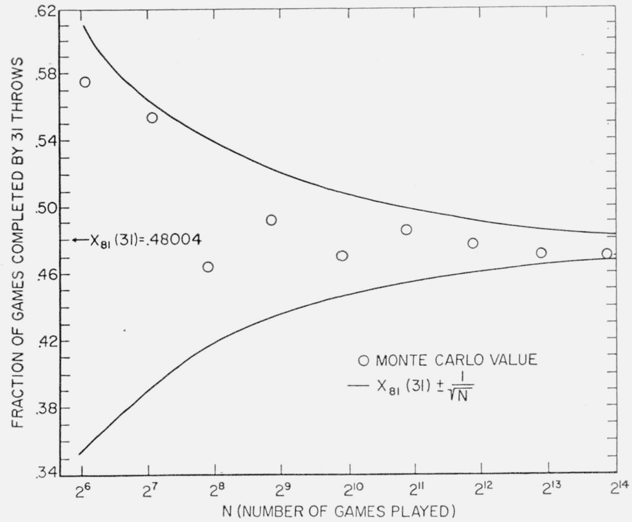

FIGURE 2. *Convergence to $X_{81}(31)$ by "$N^{-1/2}$ law."*

6. Existence of the Moments of $f(n)$

In the previous section it was assumed that the moments $\sum_{n=1}^{\infty}n^k[X_{81}(n)-X_{81}(n-1)]$ exist for all $k\geq 1$. If we expand $X(n)$ in terms of the characteristic values of the matrix P (for example see Montroll,[5] p. 419) it follows that these moments will exist if P has $\lambda=1$ as the only eigenvalue with $|\lambda|=1$. However, our Markov chain has one absorbing state which can be reached from any other state. The probability of reaching this absorbing state is independent of the initial state; but, this is a necessary and sufficient condition (see Gantmacher,[6] p. 93) that

[4] W. Feller, An introduction to probability theory and its applications (John Wiley & Sons, New York, N.Y., 1957).

[5] E. Montroll, Markov chains, Weiner integrals, and quantum theory, Comm. Pure Appl. Math. 5, 419 (1952).

[6] F. Gantmacher, The theory of matrices, II (Chelsea Press, New York, N.Y., 1959).

there exist only one eigenvalue with magnitude 1. Thus, the moments exist; also, the average length of the game can be defined by eq (3.6).

7. Concluding Remarks

The Monte Carlo calculations are in agreement with the theoretical results obtained from the matrix calculations, and the convergence according to the "$N^{-1/2}$" laws is evident from figures 1 and 2. It is worth noting that the "$N^{-1/2}$" convergence implies that if an additional significant figure is desired in a Monte Carlo calculation the computing machine time requirements are increased by a factor of 100.

(Paper 64B4–36)