SIMULATION OF RAYLEIGH—PEARSON RANDOM WALK ON COMPUTERS

BY SYAMAL KUMAR SEN

(Central Instruments and Services Laboratory, Indian Institute of Science, Bangalore-12)

[Received on August 11, 1966]

ABSTRACT

A mathematical model for the Brownian motion of a colloidal particle, as might be observed under microscope, has been attempted on the basis of the values of random numbers and random angles of the particle movement, generated laboriously using a binary fixed point Computer HEC-2M.

INTRODUCTION

Before a problem in Statistical Physics is solved theoretically, a simple idealized mathematical model of the processes has to be made first and then the problem is attempted for solution through the usual Kinetic Theory or Statistical Mechanical Methods. It frequently happens that though the micro-processes are pictured correctly, the ultimate results become difficult to be deduced from theory because of the complex nature of the stochastic processes involved. In many complicated processes even the correct equations are difficult to derive, and very gross approximations are to be made marring the original micro-picture to a degree of almost obscurity. The results so obtained through a gross distortion of the stochastic processes very often fail to agree with the observed results as it should be expected. But the stochastic processes so visualised may be entirely correct, only difficulty being the inadequacy of the mathematical tool.

All stochastic processes are but certain random processes. The elementary physical processes are usually very simple and can be expressed in very simple terms and by very simple mathematical expressions involving some type of random numbers. The stochastic processes involving neutron generation in an atomic reactor are very simple. The elementary processes of neutron generation and transport can be described mathematically by simple expressions, but to obtain the statistical behaviour of the whole reactor through the usual mathematical methods is very complicated. Monte Carlo[7] methods are extensively used for the solution using fast digital computers. Such methods are really some sort of simulation of the elementary processes on the Computer. Giving sufficient time a Computer can simulate a large number of the elementary processes from which a statistical picture of the whole thing can be obtained. At least a Computer can generate a sample of sufficient magnitude to get an estimate of the whole situation by employing the usual Statistical methods. A very well-known physical phenomenon of Brownian motion[9] of Colloidal particles has been chosen here for study. The Kinetic Theory of Brownian motion and Diffusion of Colloidal particles have long been made classic through the pioneering work of Albert Einstein and experimental confirmation of his theory has long ago been made by Jean Perrin, Svedberg and others. The simple stochastic picture of the processes was formulated by Lord Rayleigh[5] and a complete solution of the problem was due to Kluyver.

In this work the simple Rayleigh model has been chosen for convenience. The picture of the Brownian motion of a Colloidal particle is imagined to be that of a man who walks a fixed distance and then turns abruptly at any random angle from his traversed path and then walks the fixed distance and turns again. The probability of finding the man at a certain distance from his starting point is the problem. A large number of such random walks can be generated on Computer, which will also generate the necessary random angles. When a sufficient sample of such 'random walks' is generated, the results can be tested by Statistical Methods with the theoretical solution, which is regorous. This has been done in the following pages. Two graphs of the Brownian motion generated by the Computer are attached so as to give a physical picture of the Brownian motion. The pictures are very similar to those obtained by Perrin and others by camera Lucida or by long exposure photographs.

STATEMENT OF THE PROBLEM

The problem as propounded by Pearson[2] (in the case of two dimensional displacements) was as follows :

A man starts from a point 'O' and walks a distance 'a' in a straight line, he then turns through any angle whatever and walks a distance 'a' in the second straight line. He repeats the process 'n' times.

For different walks* from the same fixed point 'O' and for fixed number of steps, the 'different distances' (which we call the r-sample) are to be found out. Their x and y components are also to be computed. The r-sample is to be tested to be Maxwellian, x and y-samples to be Normal. The computed Sample mean and Standard deviations are to be compared with the Population mean and Standard deviations.

THEORY

Let the probability that the distance of the man from 'O' will lie between 'O' and a distance 'r' after 'n' steps be $P_n\ (r, a\ldots a/2)$. Then for a two-dimensional random walk

---

* Here one walk=25 steps ; it begins every time from '0' and ends after 25th step.

$$P_n(r, a \ldots a/2) \sim \left(\frac{r^2}{na^2}\right) {}_1F_1\left(1, 2, -\frac{r^2}{na^2}\right)$$

where

$${}_1F_1\left(1, 2, -\frac{r^2}{na^2}\right) = \frac{\exp[-(r^2/na^2)] - 1}{-(r^2/na^2)}$$

represents the Hypergeometric function. On simplification,

$$P_n(r, a \ldots a/2) = 1 - \exp\left(-\frac{r^2}{na^2}\right),$$

for large n.

The probability that the position of the man after 'n' steps will lie between r and $r + \delta r$ is

$$f(r) = \frac{dP_n}{dr} = \frac{2r}{na^2} e^{-r^2/(na^2)} \tag{i}$$

Since the x and y components of r are independent, we have, by applying Theorem of Compound Probability, the probability density function for x is

$$f(x) = \frac{\sqrt{(2h)}}{\sqrt{(2\pi)}} e^{-hx^2} \tag{ii}$$

and for y is

$$f(y) = \frac{\sqrt{(2h)}}{\sqrt{(2\pi)}} e^{-hy^2} \tag{iii}$$

where

$$h = \frac{1}{na^2}.$$

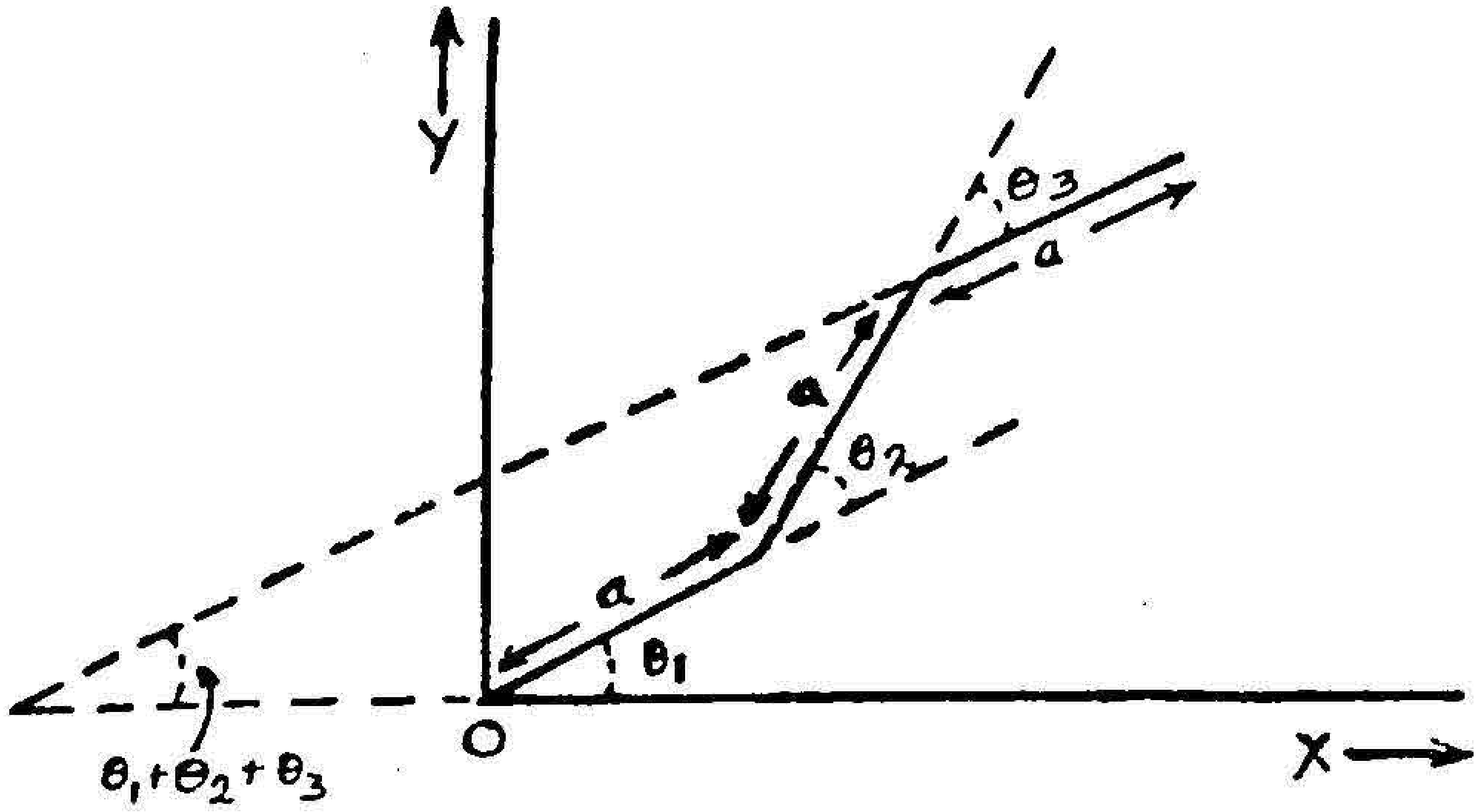
Now the distributions (ii) and (iii) are Normal in character. From (ii) and (iii) Population mean for both x and y samples is zero and

$$\sigma_x^2 = \frac{na^2}{2} \quad \text{and} \quad \sigma_y^2 = \frac{na^2}{2}.$$

From (i), Population mean for r-sample is

$$\frac{\sqrt{(n\pi)}}{2} \quad \text{and} \quad \sigma_r^2 = na^2\left(1 - \frac{\pi}{4}\right)^2.$$

Method



The heavily marked line shows only the three steps of a single ramdom walk. The x component of the 1st step is $x_1 = a\cos\theta_1$, of the first and second steps combined is $x_2 = x_1 + a\cos(\theta_1 + \theta_2)$. Generally for '$n$' steps, $x_n = x_{n-1} + a\cos(\theta_1 + \theta_2 + \theta_3 + \ldots\ldots + \theta_n)$. Similarly $y_n = y_{n-1} + a\sin(\theta_1 + \theta_2 + \theta_3 + \ldots\ldots + \theta_n)$, where $\theta_1, \theta_2, \ldots\ldots, \theta_n$ are the first, second, ..., nth random angles.

Now if r_n is the distance of the man from 'O' after n steps, $r_n = \sqrt{x_n^2 + y_n^2}$

PROGRAMMING ASPECTS

114 *Random Numbers* were selected from the *Rand Corporation* Tables of Random Numbers. Let one *Random Number* be X (converted to 18 bit number). Then X^2, a 36 bit number, was found out. Since the middle two bits were the most random bits as tested by Computer, the 18th bit was taken and was stored at the 1st bit position of a desired '10 bit random Number'. Now X was replaced by the middle 18 bits of X^2 and the process was repeated ten times to get the full '10 bit random Number'. The desired '10 bit random Number' thus formed was brought to scale zero in Binary Computer *i.e.*, $0 \leq$ 10 bit random Number < 1.

Now the 1st random angle θ_1 was $2\pi \times 10$ bit random Number, where $0 \leq \theta_1 < 2\pi$. Therefore $x_1 = a\cos\theta_1$ and $y_1 = a\sin\theta_1$ were found out where 'a' was taken as unity, r_1 was also found out, as $r_1 = \sqrt{x_1^2 + y_1^2}$.

From the contents of X, the second random angle θ_2 was found out by exactly the same procedure as above and consequently x_2, y_2 and r_2 were found out.

Thus for one *Random Number*, x_i, y_i and r_i, where $i = 1, 2, \ldots, 25$, were computed. Then we took the second *Random Number* and treated it in the same way as the first one and so on.

RESULTS

For 114 *Random Numbers*, we had 114 different walks. In each walk the man went through 25 steps and for each step we had one value of x, one value of y and one value of r. Here computed value of χ^2 for n (number of steps) = 5,10,15,20 and 25 had been shown and Tables showing the calculation of χ^2 had been given for x_{10} and r_{10} samples only.

TABLE I

Shows the computed value of χ^2 and actual value of χ^2 at 5% level.

Sample : x_{10}, $n = 10$, Computed Standard Deviation (C.S.D.) = 2.13, Sample mean (S.M.) = − 0.08

The spectrum $(-\infty, +\infty)$ of the parent variable x_{10} was divided into 10 intervals $(-\infty, -4)$, $(-4, -3)$, . . . $(4, \infty)$.

Class limit a_{10j}	Std. Class limit z_{10j}	Diff. in area under Normal curve $(p_{10})_{oj}$	Expected frequency $m'(p_{10})_{oj}$	Obs. freq. $(\nu_{10})_{oj}$	$\frac{\overline{(\nu_{10})_{oj}-m'(p_{10})_{oj}}^2}{m'(p_{10})_{oj}}$: p	Computed value of χ^2 : Σp	Actual value of χ^2 at 5% level : Q	Degrees of freedom d.f.
$-\infty$	$-\infty$							
−4	−1.84	.0319	3.64	4	0.0356			
−3	−1·37	.0524	5.97	7	0.0178			
−2	−0.90	.0988	11.26	9	0.4536			
−1	−0.43	.1495	17.04	15	0.2442			
0	0.04	.1823	20.78	27	1.8618	4.8730	14.07	7
1	0.51	.1791	20.42	20	0.0086			
2	0.98	.1415	16.13	13	0.6074			
3	1.45	.0900	10.26	9	0.1547			
4	1.92	.0461	5.26	5	0.0129			
∞	∞	.0264	3.01	5	1.3156			

TABLE II

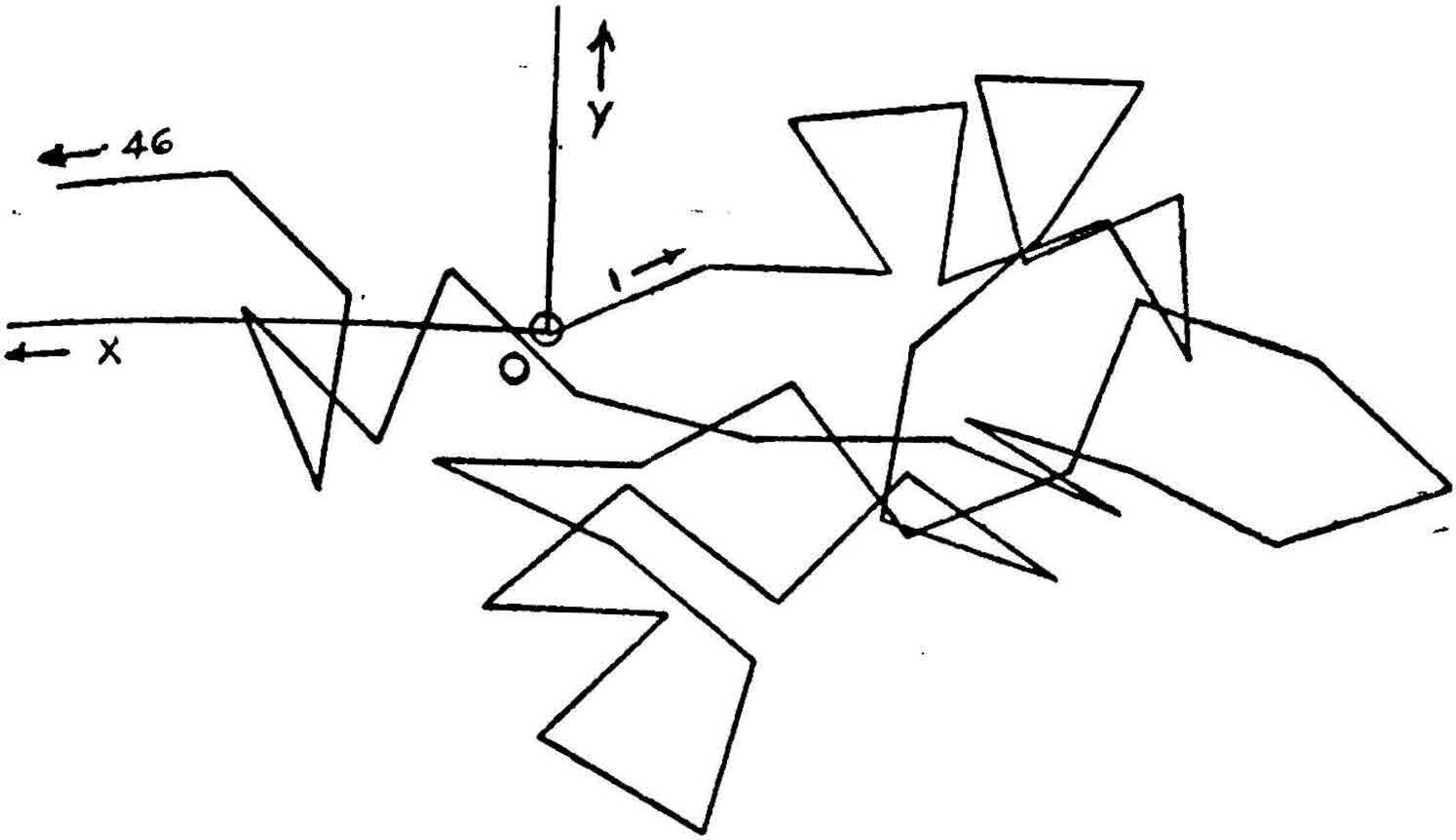
Sample : r_{10}, $n = 10$, C.S.D = 1.36, S.M. = 2.6842, No. of intervals = 7

Class limit	$P_n = 1 - \exp\left(-\frac{r^2}{na^2}\right)$	Expected freq. $m'(p'_{10})_{oj}$	Obs. freq. $(r_{10})_j$	$\frac{\overline{(v_{10})_j - m'(p'_{10})_{oj}}^2}{m'(p'_{10})_{oj}}$: p	Σ p	Q	d.f.
0	.0000						
1	.0951	10.84	13	.4304			
2	.3297	26.74	23	.5231			
3	.5935	30.07	35	.8083			
4	.7981	23.32	23	.0044	2.7352	14.07	7
5	.9179	13.66	14	.0085			
6	.9727	6.25	5	.2500			
7	.9926	2.27	1	.7105			

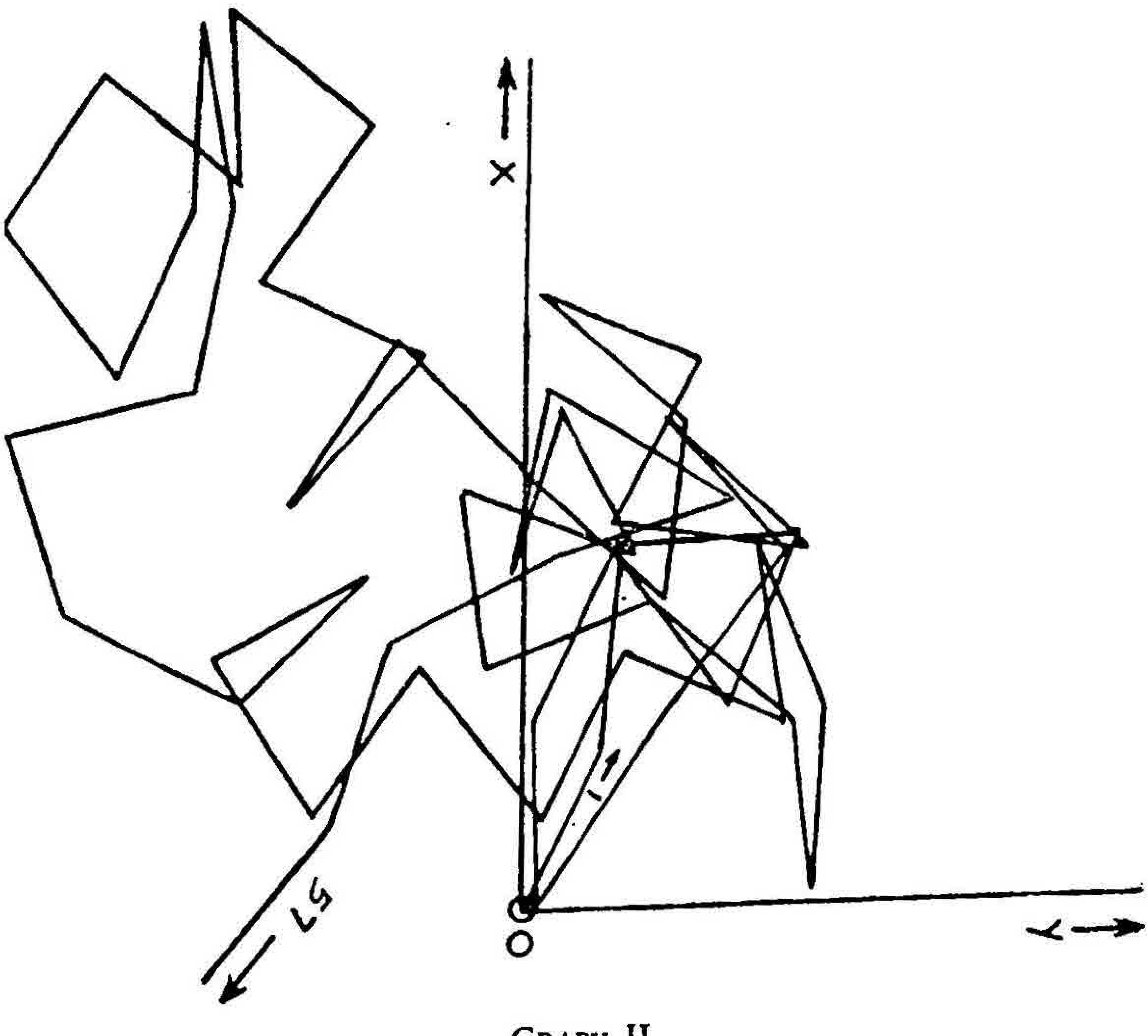
TABLE III

Shows a comparative study between Sample mean (S.M.) and Population mean (P.M.), between Computed Standard Deviation (C S.D.) and Theoretical Standard Deviation (T.S.D:) and between computed value of $\chi^2(\Sigma p)$ and actual value of $\chi^2(Q)$ at 5% level.

Sample	S.M.	P.M.	C.S.D.	T.S.D.	n	d.f.	p	Q
x_5	.11	0	1.50	1.58	5	4	7.47	9.49
x_{10}	−.08	0	2.13	2.24	10	7	4 87	14.07
x_{15}	−.35	0	2.48	2.74	15	9	3.08	16 92
x_{20}	−.13	0	2.91	3.16	20	9	4.41	16.92
x_{25}	.16	0	3.36	3.54	25	9	11.93	19.92
y_5	.16	0	1.42	1.58	5	4	2.67	9.49
y_{10}	.43	0	2.12	2.24	10	7	5.44	14.07
y_{15}	.52	0	2.56	2.74	15	7	6.86	14.07
y_{20}	.37	0	2.84	3.16	20	9	9.93	16 92
y_{25}	.41	0	3 40	3.54	25	9	8.31	16.92
r_5	1.96	1.98	1.06	1.03	5	5	2.34	11.07
r_{10}	2.68	2.80	1.36	1.47	10	7	2.74	14.07
r_{15}	3.14	3.43	1.74	1.79	15	8	4.35	15.51
r_{20}	3.74	3.96	2 07	2 07	20	9	5.73	16.92
r_{25}	4.20	4.43	2.22	2.32	25	10	3.90	18.31



GRAPH I



GRAPH II

Graphs I and II show the actual path traversed by the *Man* from the fixed Starting point 'O' for a single random *Walk*

DISCUSSION

In TABLE III, the computed value of $\chi^2 (\Sigma p)$ in the case of x and y-samples lies outside the critical region (Q, ∞). Therefore the hypothesis that all the distributions for x and y-samples are Normal in character, might be accepted at 5% level. And in the case of r-sample Σp lies outside the critical region (Q, ∞). Therefore the hypothesis that the distributions for r-sample are Maxwellian in character, might be accepted at 5% level.

The Sample means and Standard Deviations were compared with the Population means and Standard Deviations. This showed that the computations were quite satisfactory so far as the total walks (114 random walks only) were concerned.

Graphs showed the actual path traversed by the man from 'O' for a single random walk. It may be to notice that with increased number of steps the man would ultimately go away from the starting point 'O', that is to say that the Colloidal particle inside a vessel would ultimately move away from the centre of the vessel and would strike the wall causing the pressure on the wall of the vessel.

REMARK

The generated random angle should assume any value from 0 to 2π including 0 and 2π; but as the generated '10 bit random Number' in binary scale zero was always less than 1, the generated random angle could not be 2π at all. That was why the error in the length of the steps might increase as we proceeded steps higher and higher. This might impose appreciable deviation in results in larger n.

This could be improved by generating a random Number of larger number of bits, but that time it would increase machine time. So a considerable compromise was made between the machine time and number of bits, so that the randomness of the generated angle might be satisfactory.

So far as the machine time was concerned, each single random walk calculation required 12.5 minutes in an automatic electronic digital Computer HEC2M. Moreover due to fluctuation of voltage or someother reasons the machine might go wrong or might jump to some other instruction during the long time run. That was why results should be printed twice to see any difference. If no difference was detected, the result might be granted as correct. For this purpose, machine time became doubled.

CONCLUSION

The main interest of this problem is to show the randomness of the walks, as the walks are random, the distribution of a large number of walks for a particular number of steps in every walk, follows the Maxwell's distribution law for velocities. The second interest is to form a crude Mathematical model for Brownian motion of a Colloidal particle observed under Microscope.

ACKNOWLEDGMENT

The author wishes to express his sincere thanks to Prof. S. K. Mitra of Indian Statistical Institute, Calcutta for his able guidance and to Prof. P. L. Bhatnagar of Indian Institute of Science, Bangalore for taking keen interest in the preparation of this paper and for fruitful discussions.

REFERENCES

1. Watson .. .. *A treaties on Bessel's function*, 2nd edition, 1944, Art. 13-48, 419.
2. Pearson K. .. .. *Nature*, 1905, 72, 294, 342.
3. —— .. .. Drapers' Company Research Memoirs, Biometric Series, III, 1966.
4. —— .. .. *Proc. Section of Sci., K. Akad. Van Wet te Amsterdam*, 1906, 8, 341.
5. Rayleigh .. .. *Phil. Mag.* (6), 1919, 37. 321.
6. .. .. Scientific Papers, 1920, 6, 604.
7. Jaussky, O. and Todd J. .. Generation and testing of Pseudo Random Numbers, Symposium on Monte Carlo Methods, March 1954, 15.
8. Ming Chen Wang and Uhlenbeck, G. E. *Reviews of Modern Physics*, 1945, 17, 323.
9. Doob, J. L. .. .. *Annals of Maths.*, 1942, 43, 351.
10. Chandrasekhar, S, ... *Reviews of Modern Physics*, 1943, 15, 1.
11. Uhlenbeck, G. E. & Ornstein, L. S. *Physical Reviews*, 1930, 36, 823.
12. Kac, M. .. .. Am. Math. Monthly, 1946, 54, 369.
13. Rayleigh .. .. Nature, 1905, 72, 318.