Journal of Research of the National Bureau of Standards Vol. 60, No. 4, April 1958 Research Paper 2847

Radial Distribution of the Center of Gravity of Random Points on a Unit Circle [1]

F. Scheid [2]

This paper describes some Monte Carlo computations carried out on the Standards Electronic Automatic Computer (SEAC) which are related to the problem of random walks.

In 1905 Pearson [1] [3] suggested the problem of a random walk involving n steps of equal length and arbitrary direction in the plane. In 1906 Kluyver [2] obtained the solution

$$P(r,n)=r\int_0^\infty [J_0(x)]^n J_1(rx)dx$$

giving the probability of being no farther than r from the starting point after n steps of unit length. This integral has been calculated by Greenwood and Durand [3] for $n=6(1)24$, and $r=0.5(.5)n$, using punched-card tables of the Bessel functions. These tables were not extensive enough to handle the values $n=3,4$, and 5. For $n=2$ one finds by a simple argument, $P(r,2)=(2/\pi)$ arcsin $(r/2)$.

The problem described in the title is essentially the same as Pearson's. Its solution for $n=3,4$, and 5 has been obtained to roughly three decimal places by a direct sampling procedure, using SEAC. A set of n pseudorandom numbers is generated by using the multiplication method of Taussky and Todd [4]. A linear transformation converts these into a sample from the uniform distribution over $(-\pi,\pi)$. The center of gravity is found and classified in one of the intervals $j/32\leq R<(j+1)/32$, $j=0, \ldots, 31$. A new sample is then taken. Both frequency table and cumulative distribution are printed. Computation of Kluyver's integral for these values of n is possible but lengthy. If the integrand is machine computed as needed, many hours of computer time are consumed because of the two independent parameters r and n. For limited accuracy the sampling approach seems better. The distributions for $n=2$ and 6 have also been obtained for accuracy checks. Using a theorem of Kolmogoroff [5], a brief calculation indicates that, for $n=2$, the probability of a maximum error less than or equal to 0.01 after 6,000 samples have been taken is approximately one-half. A glance at table 1 shows that in fact the

[1] Preparation of this paper was initiated while the author was a member of the Numerical Analysis Training Program held at the National Bureau of Standards and sponsored by the National Science Foundation, 1957.

[2] Present address: Boston University, Boston, Mass.

[3] Figures in brackets indicate the literature references at the end of this paper.

TABLE 1.

32R	n=2 Freq	n=2 Cum	n=2 Exact	n=3 Freq	n=3 Cum	n=4 Freq	n=4 Cum	n=5 Freq	n=5 Cum	n=6 Cum	n=6 G–D
1	121	.0197	.0199	7	.001	36	.005	28	.004	.0000	.0000
2	133	.0413	.0398	37	.007	87	.018	80	.016		
3	126	.0618	.0598	58	.017	128	.038	158	.040		
4	124	.0820	.0798	67	.028	169	.063	185	.068		
5	129	.1030	.0999	95	.043	209	.094	210	.099		
6	111	.1211	.1201	113	.061	192	.123	308	.146		
7	123	.1411	.1404	141	.084	266	.163	381	.203		
8	115	.1598	.1609	172	.112	289	.207	361	.257	.2910	.2932
9	129	.1808	.1816	224	.149	238	.242	382	.314		
10	142	.2039	.2023	336	.203	316	.290	350	.367		
11	123	.2240	.2234	466	.279	335	.340	361	.421		
12	138	.2464	.2447	344	.335	360	.394	384	.479		
13	126	.2669	.2663	291	.383	357	.448	358	.533		
14	157	.2925	.2883	285	.429	365	.503	384	.590		
15	126	.3130	.3106	269	.473	365	.558	374	.647		
16	125	.3333	.3333	255	.514	405	.618	330	.696	.7665	.7672
17	150	.3577	.3565	223	.551	353	.672	317	.744		
18	158	.3835	.3803	189	.581	255	.710	324	.802		
19	135	.4054	.4047	208	.615	275	.751	273	.834		
20	148	.4295	.4298	185	.645	262	.790	224	.867		
21	157	.4551	.4558	215	.680	182	.818	179	.894		
22	158	.4808	.4826	197	.712	159	.842	143	.916		
23	173	.5090	.5106	183	.742	163	.866	131	.935		
24	190	.5399	.5399	201	.775	168	.892	109	.952	.9749	.9732
25	191	.5710	.5708	188	.805	167	.917	93	.966		
26	211	.6053	.6038	183	.835	131	.936	70	.967		
27	197	.6374	.6393	163	.862	102	.952	45	.983		
28	247	.6776	.6783	176	.890	87	.965	39	.989		
29	262	.7202	.7221	170	.918	87	.978	36	.994		
30	308	.7703	.7737	162	.944	76	.989	24	.998		
31	424	.8394	.8407	163	.971	45	.996	12	.999		
32	987	1.0000	1.0000	178	1.000	27	1.000	3	1.000	1.0000	1.0000
Number of samples	6144			6144		6656		6656		6656	

Table 2.

32R	$n=3$ Freq	$n=3$ Cum	$n=4$ Freq	$n=4$ Cum	$n=5$ Freq	$n=5$ Cum
1	4	.001	9	.002	7	.002
2	10	.003	29	.007	17	.006
3	17	.006	50	.017	21	.011
4	19	.010	59	.029	40	.021
5	48	.019	71	.043	61	.036
6	45	.028	88	.060	72	.053
7	76	.043	69	.073	99	.077
8	92	.061	109	.095	113	.105
9	107	.082	133	.121	106	.131
10	160	.113	126	.145	120	.160
11	289	.169	151	.175	142	.195
12	203	.209	191	.212	149	.231
13	203	.249	221	.255	177	.274
14	176	.283	231	.300	167	.215
15	177	.318	232	.346	227	.371
16	185	.354	287	.402	217	.424
17	170	.387	248	.450	242	.483
18	172	.421	237	.496	234	.540
19	176	.455	235	.542	237	.598
20	196	.493	235	.588	207	.648
21	185	.529	215	.630	208	.699
22	196	.568	191	.667	178	.742
23	196	.606	229	.712	179	.786
24	172	.639	210	.753	168	.827
25	184	.675	205	.793	148	.863
26	224	.719	183	.829	156	.901
27	220	.762	188	.866	117	.930
28	236	.808	183	.901	97	.954
29	225	.852	167	.934	77	.972
30	239	.899	152	.964	65	.988
31	239	.946	120	.987	34	.997
32	279	1.000	66	1.000	14	1.000
Number of samples	5120		5120		4096	

Washington, September 16, 1957.

maximum error is 0.004 for this distribution. Since the inverse square root law is operative, to obtain four decimal-place accuracy would require perhaps two-thousand times as many samples, or a year of SEAC time.

Parallel results are given in table 2 for samples drawn from a triangular distribution over $(-\pi,\pi)$. The pseudorandom numbers are modified by a direct method [6] to produce the required sample.

The question how best to compute a distribution function is not, of course, answered in this paper. It is suggested, however, that even in cases where the solution can be written explicitly, circumstances may make computation by direct sampling preferable.

References

[1] K. Pearson, The problem of the random walk, Nature (1905).

[2] J. C. Kluyver, A local probability problem, Nederl. Akad. Wetensch., Proc. (1906).

[3] Greenwood and Durand, The distribution of length and components of the sum of n random unit vectors, Ann. Math. Statistics (1955).

[4] O. Taussky and J. Todd, Generation and testing of pseudo-random numbers, Symposium on Monte Carlo Methods (Univ. of Florida, 1954).

[5] W. Feller, On the Kolmogorov-Smirnov limit theorems for empirical distributions, Ann. Math. Statistics (1948).

[6] J. von Neumann, Various techniques used in connection with random digits, NBS AMS 12 (1951).