Statistics. — *A simple technique for producing random sampling numbers.*
By H. C. HAMAKER. (Laboratorium voor Wetenschappelijk Onderzoek der N.V. Philips' Gloeilampenfabrieken, Eindhoven, Netherlands.) (Communicated by Prof. H. B. G. CASIMIR.)

(Communicated at the meeting of January 29, 1949.)

Summary.

This note describes a simple dice-throwing technique for producing random sampling numbers, together with the results of tests applied to prove the randomness of the procedure.

Having in war time no access to existing tables, I made some experiments to produce random sampling numbers for my own use which have led to the technique described below. The various tests as described by KENDALL and BABINGTON SMITH [1]), when applied to series of from 10,000 to 40,000 throws, did not show any signs of bias, so that it may be concluded that the procedure is a truly random one. Besides, the technique is extremely simple and may easily be used by anyone wishing to construct his own tables; it might, for instance, be useful as an exercise in the training of students.

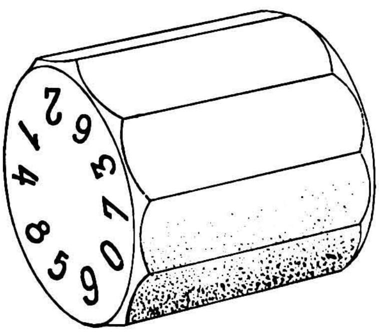

Fig. 1. The ten-sided die. Length 30 mm, diameter 30 mm.

The only tool needed is a ten-sided die as shown in fig. 1. My own dice were made of brass, but they should preferably be made of some lighter metal. The ten sides have been marked from 0 to 9 in random order by numbers engrafted in one of the top faces, as shown in the figure.

In a first series of experiments the die was thrown across the table and the uppermost number noted after it had come to rest, but bias was evident already after 1000 throws.

It was noted that the surface of the table was not perfectly horizontal, so that when rolling down-hill the die would find its final position more hesitatingly than when rolling up-hill. As this might have some influence, separate series of throws were carried out in either direction with the results recorded in table I.

TABLE I.

Frequencies observed in 600 throws, up-hill and down-hill.

Digit	Down-hill		Up-hill	
	Die No 1	Die No 2	Die No 1	Die No 2
0	86	19	48	46
1	42	41	63	58
2	50	56	65	72
3	42	85	64	65
4	34	26	63	54
5	37	57	48	51
6	73	66	72	56
7	113	65	61	70
8	51	68	55	54
9	72	117	61	74
Total	600	600	600	600
χ^2	97.2	120.4	8.6	13.90
ν	9	9	9	9
P	$\sim 10^{-15}$	$\sim 10^{-21}$	0.473	0.126

Two different dice were used, which both exhibit a pronounced bias in a set of 600 down-hill throws. By plotting the frequencies of the digits in the order in which they occur along the circumference of the dice, we obtain the two curves shown in fig. 2: die No 1 gives a curve with a single

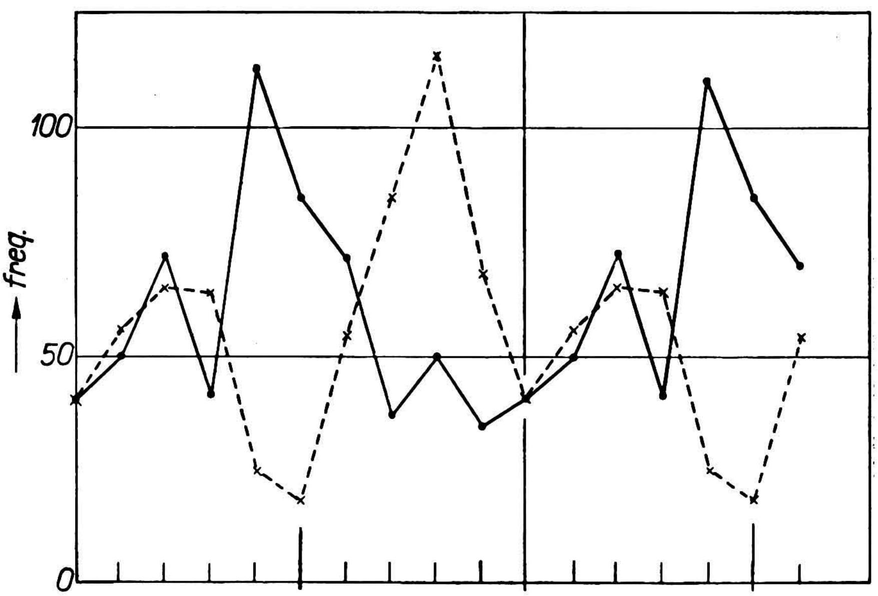

Fig. 2. The frequencies of single digits in 600 down-hill throws in consecutive order.

maximum, which might be due to excentricity of the centre of gravity, but die No 2 shows a curve with two maxima, and this cannot be so explained.

A die rolling down hill loses energy by friction and gains energy by loss of height, the total being a loss since the die ultimately comes to rest. If the amount of energy lost in each step becomes very small, slight differences in the size of the faces of the die, or in the shape of the edges, may have a pronounced influence. I think it likely that the bias observed in the down-hill throws must be explained in this way.

In keeping with this interpretation the bias in the up-hill throws, though perhaps still present, is so much reduced that it is not yet apparent in 600 throws: the die now loses energy both by friction and by a gain in height, and its final position will mainly be determined by its initial energy.

It may be added that the slope of the surface was only $\frac{1}{2}^{\circ}$, a slope of 1° being already sufficient to keep the die rolling down-hill when once set in motion.

It was inferred from these observations that satisfactory results might probably be obtained, if a die in spinning motion is suddenly stopped by a strong frictional force, and this principle was brought in practice in the following simple way.

From the flat hand a die (as shown in fig. 1) was thrown spinning into the air and caught again in its downward fall; one of the ten side-faces was then selected by the thumb and the corresponding digit read off. To preclude personal bias the top face with the digits on it should not be visible while the choice is being made, a measure that can easily be effected.

This practice proved entirely satisfactory. By throwing with the left hand and recording the digits on a typewriter with the right, 1000 throws could be completed in 40 minutes. At a later stage two dice were thrown simultaneously, one with each hand, an assistent recording the scores; in that way 1000 throws took not more than half an hour, including the insertion of a new sheet in the typewriter and a few corrections for misprints.

As it does not seem likely that in this throwing technique one particular digit should have definite preference above the others, a correlation between successive throws must be considered as the main form of bias conceivable. And a bias of this kind is not very probably either, since both the throwing and the catching of the die are operations in which many random factors play a role, and in which it is not easy to achieve a great degree of regularity even after prolonged practice; it was estimated that a die made, on the average, about 5 revolutions per throw.

But the final decision as to the absence or presence of bias should rest with the scores obtained. To settle this question the frequencies of "overlapping" pairs have been recorded in a set of 10,000 singlehanded throws; that is, a series of scores

61 80 38 05

was recorded as 61, 18, 80, 03, 38, 80, 05, and so on. Thus each single throw is recorded twice; once as the second and once as the first digit in a pair. Hence the marginal totals of rows and of columns in the 10 × 10 frequency array (table II) are equal to each other and equal to the total

TABLE II.

Frequencies of 'overlapping' pairs in a series of 10,000 throws.

Second digit / First digit	0	1	2	3	4	5	6	7	8	9	Total
	Frequencies										
0	112	103	98	99	105	98	91	98	96	101	1,001
1	75	100	91	95	93	108	107	103	100	105	977
2	107	87	105	93	106	100	103	91	105	96	993
3	87	105	96	82	88	103	106	100	115	84	966
4	100	102	103	86	111	102	104	73	103	112	996
5	112	101	105	105	102	100	98	95	100	99	1,017
6	104	108	85	117	102	94	99	97	104	99	1,009
7	102	83	98	95	86	112	93	103	102	109	983
8	103	97	104	101	96	95	115	121	94	105	1,031
9	99	91	108	93	107	105	93	102	112	117	1,027
Total	1,001	997	993	966	996	1,017	1,009	983	1,031	1,027	10,000

$\chi^2 = 73.64$ $\quad \sqrt{2\chi^2} = 12.14$ $\quad P = 0.214$

$\nu = 90$ $\quad \sqrt{2\nu - 1} = 13.30$

$\Delta = -1.16$

TABLE III.

Frequencies of 'independent' pairs in 20,000 double-handed throws.

Second digit / First digit	0	1	2	3	4	5	6	7	8	9	Total
	Frequencies										
0	101	102	99	91	97	114	102	98	91	92	987
1	101	85	103	93	122	100	110	115	100	93	1,022
2	108	90	111	91	104	102	110	102	102	92	1,012
3	97	101	103	108	95	94	97	89	103	97	984
4	103	91	106	120	109	96	110	95	91	105	1,026
5	83	101	109	97	114	91	98	94	79	92	958
6	96	127	98	98	100	95	107	100	101	102	1,024
7	99	107	89	110	103	98	98	87	76	121	988
8	101	113	105	107	107	105	83	94	100	100	1,015
9	103	88	85	94	108	107	101	92	115	91	984
Total	992	1,005	1,008	1,009	1,059	1,002	1,016	966	958	985	10,000

$\chi^2 = 83.40$ $\quad \sqrt{2\chi^2} = 12.92$ $\quad P = 0.263$

$\nu = 99$ $\quad \sqrt{2\nu - 1} = 14.04$

$\Delta = -1.12$

frequencies with which the single digits were observed. It is easily deduced that the number of degrees of freedom for the entire array is 90.

As indicated at the bottom of the table the χ^2-test applied to the entire array gives $P = 0.214$, and when applied to the marginal totals (see table IV A) $P = 0.893$; reasonable values which do not indicate bias.

A second series of 10,000 double-handed throws gave a sequence of 20,000 random digits, and the frequencies of these arranged in 10,000 'independent' pairs have been collected in table III.

The two sets of marginal totals are now independent and their sum is equal to the frequencies of the single digits; the number of degrees of freedom for the entire array is 99. The χ^2-test yields $P = 0.263$ for the complete 10×10-array and $P = 0.527$ for the single-digit frequencies (see table IV B).

Finally in another set of 10,000 singlehanded throws the frequencies of the single digits were computed, and these were added to the sum of the corresponding frequencies in tables II and III. This gave us the single-digit frequencies for a total of 40,000 throws as shown in table IV C; $P = 0.696$, again a normal value.

TABLE IV.

Single-digit frequencies in various cases.

0	1	2	3	4	5	6	7	8	9	Total
A. In 10,000 single-handed throws (table II)										
1,001	977	993	966	996	1,017	1,009	983	1,031	1,027	10.000
$\chi^2 = 4.23$;		$\nu = 9$;		$P = 0.893$						
B. In 20,000 double-handed throws (table III)										
1,979	2,027	2,020	1,993	2,085	1,960	2,040	1,954	1,973	1,969	20,000
$\chi^2 = 7.93$;		$\nu = 9$;		$P = 0.527$						
C. In a total of 40,000 throws (including tables II and III)										
4,000	3,996	4,077	3,898	4,058	3,957	4,008	3,950	4,027	4,029	40,000
$\chi^2 = 6.42$;		$\nu = 9$;		$P = 0.696$						

Thus our data satisfy the frequency and the series tests [1]), which provides fairly conclusive evidence that the technique adopted is reliable.

In constructing their table of random sampling numbers KENDALL and BABINGTON SMITH [1]) used a disc with the numbers 0 to 9 inscribed on its circumference and rotating past a pointer; in momentaneous flashes at random intervals the number seen nearest to the pointer was recorded.

When using this method much will depend on the reaction velocity of the observer, and if he should fall short in this respect, there is a danger of personal bias being introduced into the records; in one case such a bias could actually be demonstrated from the frequencies of the separate digits. Similarly there might be some danger of a person being disinclined to read a certain number after it has repeatedly been observed, or being inclined to read a specified number if it has not occurred for a long time.

Probably the numbers thrown as described above are less subject to bias of this kind; one should have to be a die-hard falsifier to introduce personal bias once the choice has been made, and the only bias possible is that which may be introduced in the throwing, the catching, and the chosing of one of the ten faces.

To test for the kind of bias just mentioned KENDALL and BABINGTON SMITH [1]) devised two other tests: the poker- and the gap-test. For the sake of completeness these have also been applied to the final series of 20,000 double-handed throws giving *P*'s of 0.672 and 0.968 respectively. I shall not record the data in detail here; a fuller account has been published in Dutch in the journal 'Statistica' [2]).

I am indebted to Mr H. A. C. VAN DER LINDEN for his painstaking assistance in recording the throws and performing the tests.

Eindhoven, 15 December 1947.

LITERATURE.

(1) M. G. KENDALL and B. BABINGTON SMITH, Jl. Roy. Stat. Soc. **101**, 147 (1938); Jl. Roy. Stat. Soc. Suppl. **6**, 51 (1939).

(2) H. C. HAMAKER, Statistica, **2**, 97—106 (1948).