Statistical Science
1991, Vol. 6, No. 1, 89–97

Stochastic Simulation in the Nineteenth Century

Stephen M. Stigler

Abstract. In the last quarter of the nineteenth century, three separate (but not entirely independent) papers appeared describing methods of studying complicated statistical procedures through the generation of random normal deviates. All three authors referred to problems in the smoothing of series as a motivation; all used different methods for generating the deviates. One presented itself as a method for general use and claimed to be suitable for efficient generation of large numbers of variates. The relevant works (by Erastus De Forest in 1876, by George H. Darwin in 1877, and by Francis Galton in 1890) are reproduced.

Key words and phrases: Simulation, Monte Carlo method, history of statistics, Francis Galton, dice, random number generation.

INTRODUCTION

Simulation, in the modern sense of that term, may be the oldest of the stochastic arts. Long before the calculus of probabilities was developed for secular purposes, lots were cast to attempt to read the minds of the gods. Indeed, the faith in the power of this device for learning by analogy was sometimes so strong that wagers were placed upon the outcomes. Simulation as a tool of science, however, is of a much more recent vintage.

The question of when to begin the history of scientific simulation depends upon what one means by the term. If by simulation we mean the use of a controlled random or pseudo-random device to illustrate a mathematical theorem, we would have to go back at least to the eighteenth century. In the mid-eighteenth century the French naturalist Buffon had a child perform a series of 2,048 sets of tosses of a coin, each time the set continuing until a head occurred (Table 1). Buffon was attempting to determine an empirical value for the St. Petersburg game, and he concluded that the value of the game was about 5, despite its infinite expected value.

In the same essay that reported this empirical illustration of a geometric distribution, Buffon also described his famous "Buffon's needle" experiment, where counts of the number of times a dropped needle intersects a system of parallel lines can be used to determine the value of $\pi = 3.1416\ \ldots$ experimentally (Buffon, 1735, 1777; Perlman and Wichura, 1975). Buffon did not give data for his needle experiment, but several people pursued this in the nineteenth century with great vigor. For example, the Swiss astronomer Rudolf Wolf performed several experiments comprising over 5,000 trials starting in 1849 (Riedwyl, 1990), and Augustus De Morgan (1915, pages 283–284) reported experiments performed in 1855 and later that gave values for π of 3.1553 (based on 3204 trials) and 3.137 (based on 600 trials). Other trials were carried out in America in 1864 (Hall, 1873). Somewhat later an Italian, Lazzarini (1902), reported performing 3,408 trials and getting a value of $\pi = 3.1415929$, an agreement so close that Coolidge (1925, pages 81–82) suspected him of "watching his step," since he estimated the chance of such an agreement without optional stopping was about 1/69. These suspicions are not decreased by the information (not cited by Coolidge or others who have indicted Lazzarini) that the year before these experiments, Lazzarini had published a long article on rational approximations to π (Lazzarini, 1901)!

There is, however, a question as to whether even under the most charitable interpretation these experiments with "Buffon's needle" should be considered as instances of simulation: they were illustrations of mathematical theory and could even

Stephen M. Stigler is Professor, Department of Statistics, University of Chicago, 5734 University Avenue, Chicago, Illinois 60637.

TABLE 1
Data from Buffon's experiment on the St. Petersburg game

Tosses (k)	Payoff (2^{k-1})	Frequency
1	1	1061
2	2	494
3	4	232
4	8	137
5	16	56
6	32	29
7	64	25
8	128	8
9	256	6
		2048

The data from Buffon's experimental attempt to place a value on the St. Petersburg game, where a coin is tossed until heads occurs and 2^{k-1} is the payoff, where k is the number of tosses. Buffon had a child play the game 2048 times. (Buffon, 1777, pages 84–85; De Morgan, 1915, vol. 1, page 282; Jorland, 1987, page 168.)

be construed as checks on the fit of theory to nature, but they were not attempts to learn about a process by analogy. If the experimental value of π had come out far from 3.1416, then the experiment would have been discarded, not the value 3.1416. Indeed, as Coolidge's observation about Lazzarini shows, the experiment would also be discarded if the experimental value was too close to 3.1416! If we insist that the purpose of the simulation be to learn about the process under study, say to learn the distribution of a complicated statistic by computing its values from artificially generated samples, then we must look to a more recent period for a less ambiguous example (although it could be argued that Buffon's coin tosses fit this definition, since his aim there was to put a value on the St. Petersburg Paradox wager).

There has been a tendency in recent years to date the use of simulation in statistics only from the early years of the twentieth century. For example, Teichroew (1965) and Irwin (1978) suggest that its earliest appearance may be in "Student's" classical investigation of the distribution of the t-statistic (Gosset, 1908a), where "Student" (William S. Gosset) generated 750 samples of size 4 by shuffling 3,000 cards labeled with anthropometric measurements on 3,000 criminals and then grouping them in 750 groups of size 4. Gosset also used the same generated samples in his investigation of the distribution of the correlation coefficient (Gosset, 1908b). (See also Muller, 1978.) However, even more sophisticated uses of the technique can be found in nineteenth century literature, and it is the purpose of this note to present three such examples from the last quarter of the nineteenth century. One of these was published in America (by Erastus L. De Forest, in 1876) and two in England (by George H. Darwin, in 1877, and by Francis Galton, in 1890).

The use of the word *simulation* in this context is somewhat anachronistic; a common definition of the period would be along the lines of that given by the *Century Dictionary* (published in several editions, from 1889):

> simulation, n. 1. The act of simulating, or feigning or counterfeiting; the false assumption of a certain appearance or character; pretense, usually for the purpose of deceiving.

The *Century* entry went on to include a quotation from *Scribner's Magazine*,

> The *simulation* of nature, as distinguished from the actual reproduction of nature, is the peculiar province of stage art [Kobbé, 1888, page 438].

The three passages to be discussed here are not the product of stage art, nor were they concocted to deceive. Rather, all three concern the artificial generation of normally distributed random numbers for the express purpose of studying the properties of complicated statistical procedures. All concern simulation in the modern sense of the word, as a modern stochastic art for the study of statistical science. All three involve the generation of half-normal variates and the separate assignment of randomly generated signs to the variates, but the three involve three different randomizing devices. De Forest drew labeled cards from a box, Darwin used a spinner, and Galton employed a set of special dice (Figure 1).

1. ERASTUS L. DE FOREST (1876)

Erastus Lyman De Forest (1834–1888) was a graduate of Yale University (class of 1854) who published a number of remarkable works in mathematical statistics between 1873 and 1885. His life and work are discussed in Stigler (1978), and several of his papers (including the one involving simulation) are reprinted in Stigler (1980). A major portion of De Forest's work was motivated by problems in the smoothing of life tables, where he studied methods of local smoothing that included optimality criteria for fit (least squares) and smoothness (the minimization of expected fourth differences). In the course of devising a test for the goodness-of-fit of his smoothed mortality curves to the crude empirical curves that were his data, he was led to consider the distribution of the quantity

$$\log\left(\frac{v}{v'}\right),$$

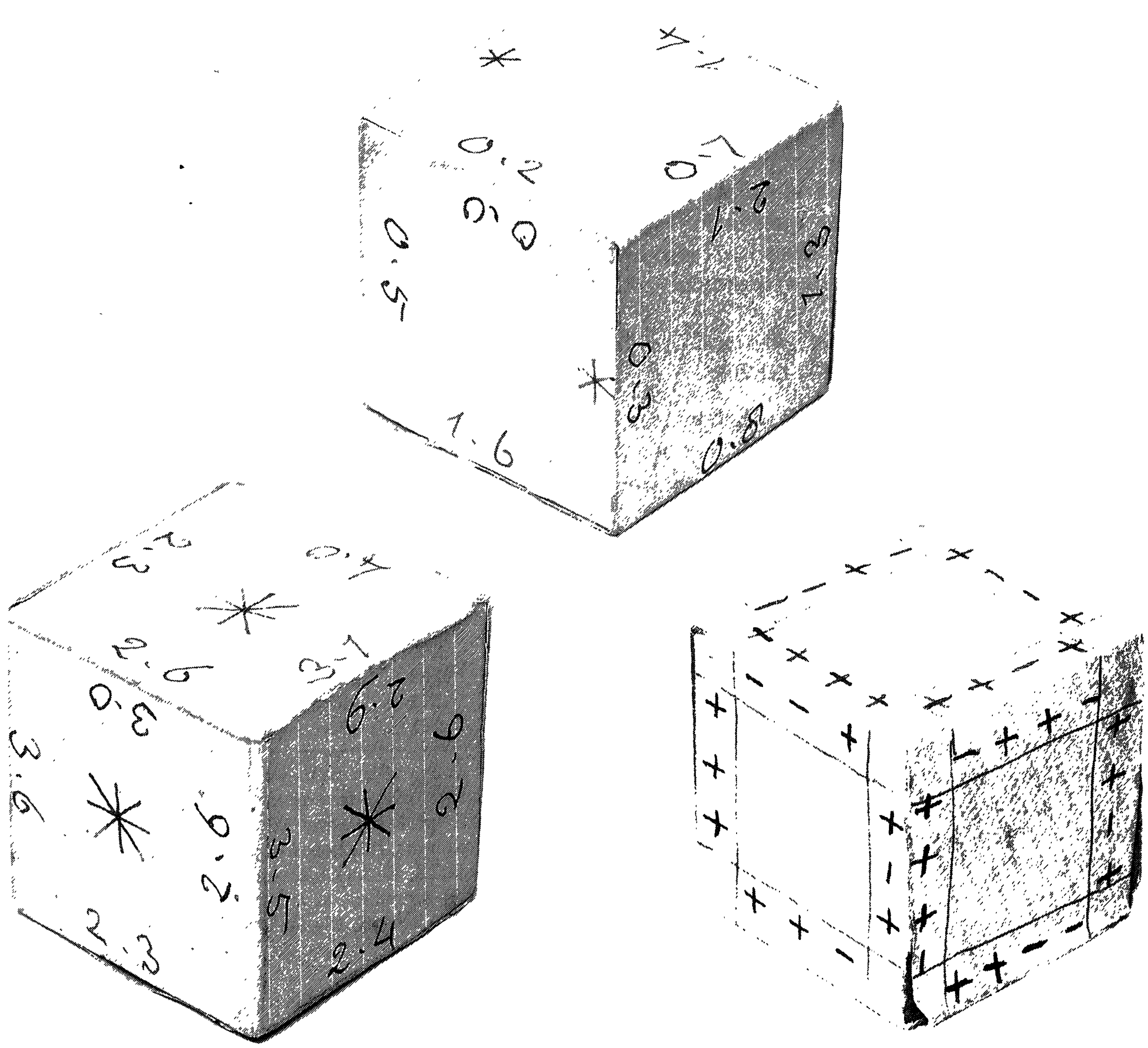


FIG. 1. *Photographs of one each of the three types of Galton's dice. These date from about* 1890 *and are perhaps the oldest surviving device for simulating normally distributed random numbers. They are presently in the Galton Collection at University College London* (*Box* 150/4).

where v and v' are the absolute values of independent standard normal random variables. (De Forest worked with logarithms to base 10.) We might now characterize the distribution of this random variable as half the logarithm of a variable with an $F(1,\ 1)$ distribution, but such a characterization was not available in 1876. Anyway, De Forest did not require the entire distribution; he would be satisfied with the "probable error" (the median deviation, or, .6745 times the standard deviation) of the arithmetic mean of m independent copies of this variable. He was also interested in the answer to this question where both v and v' are replaced by averages of n independent similar quantities. That is, De Forest required, for various m and n, the probable error E' of

$$\frac{1}{m}\sum_{i=1}^{m}\log\left(\frac{v}{v'}\right)_n = \frac{1}{m}\sum_{i=1}^{m}\log\left(\frac{\bar{v}_i}{\bar{v}'_i}\right),$$

where

$$\bar{v}_i = \frac{1}{n}\sum_{j=1}^{n} v_{ij},$$

$$\bar{v}'_i = \frac{1}{n}\sum_{j=1}^{n} v'_{ij},$$

and all v_{ij}, v'_{ij} are absolute values of independent $N(0,\ 1)$ random variables.

By a delta-method argument, De Forest found the approximation

$$(113)\qquad E' = -\sqrt{\frac{2}{m}}\log\left(1 - \frac{.6745}{\sqrt{2n}}\right).$$

It was as a check on the accuracy of this approximation that he was led to his simulation study, which we reproduce here. The simulation study, which appears to have been based on sampling without replacement, included an assessment of the accuracy of the result.

From pages 23–25 of *Interpolation and Adjustment of Series,* by E. L. De Forest (1876):

The demonstration of formula (113), however, was not a strictly rigorous one, and it has been thought desirable to test the accuracy of the formula by trials made on a sufficiently large scale, in the following manner.

The well-known function

$$P = \frac{2}{\sqrt{\pi}}\int_0^t e^{-t^2}\,dt$$

represents the probability that, in a system whose mean error is $\frac{1}{2}\sqrt{2}$, any error which occurs will not exceed t when taken without regard to sign. The function becomes zero when t is zero, and unity when t is infinite. The values of P, for all values of t taken at intervals of .01 from 0 to 2, may be found in a table appended to Vol. II. of Chauvenet's Astronomy. From that table the values of the variable t have been obtained by simple interpolation, for every value of the function P from .005 at intervals of .010 up to .995, making in all 100 values of t. These values, corresponding as they do to equal increments of .010 in the value of P, from zero up to unity, may be regarded as approximately a system of errors of equal frequency, that is, errors any one of which is as likely to occur as any other. These 100 errors are shown in the accompanying table [the present Table 2]. They have been inscribed upon 100 bits of card-board of equal size, which were shaken up in a box and all drawn out one by one, and entered in a column in the order in which they came, like the errors v' in column (5) of Table VI, if the signs are neglected. [Table VI, De Forest (1876, page 18), presented the details of an example of the calculation of his empirical estimates.]

They were again shaken up in the box and drawn out and arranged in a second column, like the errors v in column (6) of that table. From these two columns 100 values of $\log(v/v')$ were computed, just as in column (7) of the table. Also, by taking v and v' in groups of two, four and five, three other columns were formed containing the fifty, twenty-five and twenty values of $\log(v/v')_2$, $\log(v/v')_4$ and $\log(v/v')_5$ respectively, after the manner of column (8) in Table VI, making four such columns in all. Since the two systems of errors v and v' are here known to be equivalent systems, the arithmetical mean of all the values of $\log(v/v')_n$ in either of the four columns will be theoretically zero, and may be taken as such in estimating the probable error of the mean. The sum of the squares of the deviations from this mean in any column, therefore, will be simply the sum of the squares of the numbers standing in that column, so that the probable error of the arithmetical mean of all the values of $\log(v/v')_n$ in any column will be

$$E' = \frac{.6745}{m}\sqrt{\sum \log^2\left(\frac{v}{v'}\right)_n}$$

TABLE 2
Errors of equal frequency

P	t	P	t	P	t	P	t
.005	.0044	.255	.230	.505	.483	.755	.822
.015	.0133	.265	.239	.515	.494	.765	.840
.025	.0222	.275	.249	.525	.505	.775	.858
.035	.0310	.285	.258	.535	.517	.785	.877
.045	.0399	.295	.268	.545	.528	.795	.896
.055	.0488	.305	.277	.555	.540	.805	.916
.065	.0577	.315	.287	.565	.552	.815	.937
.075	.0666	.325	.296	.575	.564	.825	.959
.085	.0755	.335	.306	.585	.576	.835	.982
.095	.0844	.345	.316	.595	.589	.845	1.006
.105	.0933	.355	.326	.605	.601	.855	1.031
.115	.1023	.365	.336	.615	.614	.865	1.057
.125	.1112	.375	.346	.625	.627	.875	1.085
.135	.1202	.385	.356	.635	.641	.885	1.115
.145	.1292	.395	.366	.645	.654	.895	1.146
.155	.1382	.405	.376	.655	.668	.905	1.181
.165	.1473	.415	.386	.665	.682	.915	1.218
.175	.1564	.425	.396	.675	.696	.925	1.259
.185	.1655	.435	.407	.685	.710	.935	1.305
.195	.1746	.445	.417	.695	.725	.945	1.357
.205	.1837	.455	.428	.705	.740	.955	1.418
.215	.1929	.465	.439	.715	.756	.965	1.491
.225	.202	.475	.449	.725	.772	.975	1.585
.235	.211	.485	.460	.735	.788	.985	1.720
.245	.221	.495	.471	.745	.805	.995	1.985

This table, taken from De Forest (1876, page 24), is essentially a table of the inverse cumulative of a half-normal distribution with "probable error" (that is, median deviation) 1.0. As De Forest explains, it is based upon a table in Chauvenet (presumably Table IX, as in Chauvenet 1891, vol. 2, page 593, which gives P as a function of t). One typographical error is corrected here, where he had .234 for .235 in column one.

where m denotes the whole number of groups of n errors each. The values of E' thus obtained are of course liable to some deviation from the normal values, depending as they do on the fortuitous sequence of the errors v and v' as they come from the box. To obtain a fair average result, therefore, six separate sets of drawings were made and treated in the manner just described.* The means of the actual or observed values of E', as deduced from these six trials, are shown in the following table,

	$n = 1$ $m = 100$	$n = 2$ $m = 50$	$n = 4$ $m = 25$	$n = 5$ $m = 20$
E' (observed)	.0423	.0334	.0332	.0311
E' (theoretical)	.0398	.0357	.0335	.0329

and the theoretical values given by formula (113) are also shown for the purpose of comparison. From the differences between the separate results of the six trials and the means of them all, it appears that each of the above mean values of E' (observed) is subject to a probable error of about .0009. The agreement between observation and theory is thought to be as good as could be expected, under all the circumstances. It justifies the belief that formula (113) is accurate enough for practical purposes, furnishing a value of E' more trustworthy than could be obtained from the observed errors in any particular case, and especially so when the number of terms in the series is not large.

2. GEORGE H. DARWIN (1877)

George Howard Darwin (1845–1912), mathematician and astronomer, was the son of Charles Darwin and a cousin to Francis Galton. From 1883 he was the Plumian Professor of Astronomy and Experimental Philosophy at Cambridge, and most of his work concerned the earth, including tidal theory and dynamic meteorology. In 1877, while a Fellow at Trinity College, Cambridge, he published an article on the smoothing of series of observations and the interpolation of surfaces, with particular attention to the application to meteorological data.

Darwin's main concern was to present an "empirical rule" for smoothing series. In fact, as he discovered while the paper was in press, the local smoothing methods he discussed were similar to earlier work. A footnote stated:

> Since this paper has been in the hands of the printer, I have learnt that M. Schiapparelli has written a work entitled *Sul modo di ricavare la vera expressione delle leggi della natura dalle curve empiriche*; (Milan, 1867), and that M. De Forest has written on the subject in the 'Annual Reports of the Smithsonian Institution' for 1871 and 1873, and in the 'Analyst' (Iowa) for May 1877.

The De Forest papers he cited (De Forest, 1873, 1874, 1877) all discuss De Forest's methods for adjusting series, and in that respect there was considerable overlap (De Forest's development went further). However, Darwin made no mention of De Forest's 1876 pamphlet, which included his only treatment of simulation. Since that pamphlet was privately printed and not circulated as a periodical, and no particular emphasis was put upon simulation as a method, it is unlikely that Darwin was initially aware that De Forest had preceded him in this regard, although De Forest did cite the pamphlet in the July 1877 continuation (De Forest, 1877b) of the "Analyst" article Darwin mentioned in the footnote (De Forest, 1877a), and it is plausible that Darwin saw the work eventually. In any event, Darwin's implementation differed from De Forest's: where De Forest had drawn tickets from a box, Darwin constructed a spinner. A circular card was marked along its edge with a scale determined by a half-normal cumulative distribution. By spinning the card, Darwin could generate half-normal variates. He attached signs to these by tossing a coin, and he would then attach the resulting "errors" to a sinusoidal function in order to experiment with his "empirical rule" for smoothing series. As a test of fit, he was content to observe that "The general result of a good many trials was such as to justify the smoothing process." He did not give details on how finely the scale was graduated (or how the scale was constructed). He did advise that the disk (which he at one point referred to as "the roulette") be stopped manually while it was still spinning rapidly. Presumably this was both to diminish potential bias toward one section of the disk and to speed up the procedure.

From pages 6–7 of "On Fallible Measures of Variable Quantities," by G. H. Darwin (1877):

> The merits of an empirical rule like this must of course depend on how it seems to work practically. I therefore devised the following scheme for testing it. A circular piece of card was graduated radially, so that a graduation marked x was $720 \int_0^x e^{-x^2}\, dx/\sqrt{\pi}$ degrees

* Or, speaking more accurately, there were only four columns drawn, and these gave six different combinations of two and two [De Forest's footnote].

distant from a fixed radius. The card was made to spin round its centre close to a fixed index. It was then spun a number of times, and on stopping it the number opposite the index was read off.† From the nature of the graduation the numbers thus obtained will occur in exactly the same way as errors of observation occur in practice; but they have no signs of addition or subtraction prefixed. Then by tossing up a coin over and over again and calling heads + and tails − , the signs + or − are assigned by chance to this series of errors. About a dozen equidistant values of some function (say sine or cosine) were next taken from a Table, and the errors added to or subtracted from them in order. The errors may be made either small or large by multiplying them by any constant. The falsified values may then be fairly taken to represent a series of observations; but we here know what are the true ones. The corrections were then applied, in some cases arithmetically and in others graphically, and the deviations of the corrected values from the true were observed.

In other cases a series of equidistant ordinates were taken, and a sweeping free-hand curve was drawn to represent the true curve, and the several ordinates of this curve were falsified by the roulette and then corrected by a graphical application of the rule. The general result of a good many trials was such as to justify the smoothing process. Where the errors were considerable the mean error was much reduced, although the actual error of some ordinates was increased; where the errors were very small the mean error was even slightly increased. Although the danger of oversmoothing was obvious, and the sharpness of the features of the curve was generally diminished, yet I think it was clear that the method might generally be employed with advantage, especially in such cases as the attempt to deduce some law from statistics or a series of barometric oscillations of considerable periods. The errors must be very large to justify a quadruple operation. This method of trial could not be so well applied to testing the case of an odd number of smoothing operations, where we are left finally at intermediate ordinates.

† It is better to stop the disk when it is spinning so fast that the graduations are invisible, rather than to let it run out its course [Darwin's footnote].

3. FRANCIS GALTON (1890)

Francis Galton (1822–1911) was such a fertile source of statistical ideas over his long life that it should not cause surprise that he contributed to simulation as well. Indeed, his well-known probability machine, the Quincunx (Stigler, 1986, Chapter 8) can be viewed as a simulation device, although of a different type than those we consider here. Galton was in frequent correspondence with his cousin George Darwin, and it is possible that Darwin's idea of a simulation spinner arose from communication with Galton, although I have no direct evidence of this. (A January 12, 1877 letter from Galton to Darwin explaining the Quincunx is reproduced in Stigler (1986, pages 278–279).)

Galton's 1890 letter to *Nature*, reproduced in its entirety below, differs from the earlier works in the important respect that it purports to present a general method for simulation. The investigation of processes for smoothing or interpolation are mentioned as applications, but Galton's aim clearly goes further. And he specifically claims that his method is superior to the alternatives of shuffling cards, drawing marked balls from a bag, or spinning a roulette wheel. Galton was surely aware of Darwin's paper, and, while I doubt he had seen De Forest's, it would not have surprised him.

Galton's scheme was ingenious. For two centuries it had been apparent that a well-made die could produce a random selection from among six possibilities. Galton may have been the first to see how this could be enlarged to 24, by writing the possibilities along the edges of the die!

In essence, Galton's scheme could be viewed as a refinement of De Forest's. One die is used to provide a random selection from 24 different values. The values were taken from a table he had given in his 1889 book *Natural Inheritance* and were constructed to give a discretized version of a sample from a half-normal distribution with "probable error" (median error) 1.0. The three largest values were marked with parentheses, indicating that if they were observed, attention should instead be given to a second die with a much finer scale covering the range 2.29 to 4.55. A third die was used to attach signs to the selection; these were arranged in groups to save work.

From *Nature*, vol. 42, pages 13–14 (May 1, 1890):

DICE FOR STATISTICAL EXPERIMENTS

Every statistician wants now and then to test the practical value of some theoretical process, it may be of smoothing, or of interpolation, or of obtaining a measure of variability, or of

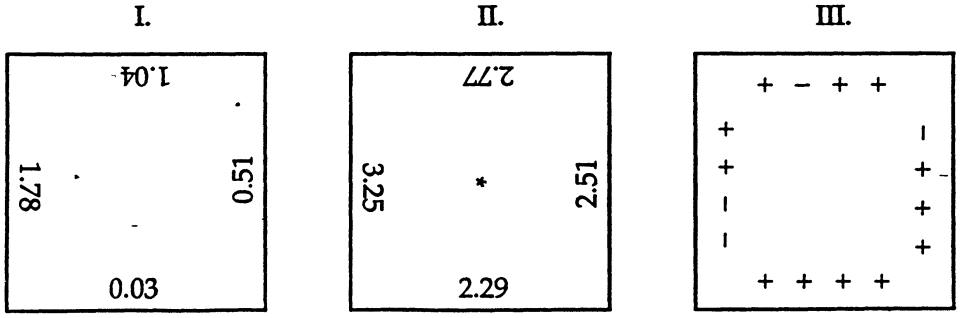


FIG. 2.

making some particular deduction or inference. It happened not long ago, while both a friend and myself were trying to find appropriate series for one of the above purposes, that the same week brought me letters from two eminent statisticians asking if I knew of any such series suitable for their own respective and separate needs. The assurance of a real demand for such things induced me to work out a method for supplying it, which I have already used frequently, and finding it to be perfectly effective, take this opportunity of putting it on record.

The desideratum is a set of values taken at random out of a series that is known to conform strictly to the law of frequency of error, the probable error of any single value in the series being also accurately known. We have (1) to procure such a series, and (2) to take random values out of it in an expeditious way.

Suppose the axis of the curve of distribution (whose ordinates at 100 equidistant divisions are given in my "Natural Inheritance," p. 205) to be divided into n equal parts, and that a column is erected on each of these, of a + or a − height as the case may be, equal to the height of the ordinate at the middle of each part. Then the values of these heights will form a series that is strictly conformable to the law of frequency when n is infinite, and closely comformable when n is fairly large. Moreover the probable error of any one of these values irrespectively of its sign, is 1.

As an instrument for selecting at random, I have found nothing superior to dice. It is most tedious to shuffle cards thoroughly between each successive draw, and the method of mixing and stirring up marked balls in a bag is more tedious still. A teetotum or some form of roulette is preferable to these, but dice are better than all. When they are shaken and tossed in a basket, they hurtle so variously against one another and against the ribs of the basket-work that they tumble wildly about, and their positions at the outset afford no perceptible clue to what they will be after even a single good shake and toss. The chances afforded by a die are more various than are commonly supposed; there are 24 equal possibilities, and not only 6, because each face has four edges that may be utilized, as I shall show.

I use cubes of wood $1\frac{1}{4}$ inch in the side, for the dice. A carpenter first planed a bar of mahogany squarely and then sawed it into the cubes. Thin white paper is pasted over them to receive the writing. I use three sorts of dice, I., II., and III., whose faces are inscribed with the figures given in the corresponding tables. Each face contains the 4 entries in the same line of the table. The diagram shows the appearance of one face of each of the 3 sorts of dice; II. is distinguished from I. by an asterisk in middle; III. is unmistakable. It must, however, be understood, that although the values are given to the second place of decimals both in the tables and in this diagram, I do not enter more than one decimal on the dice. The use of the second decimal is to make multiplication more accurate, when a series is wanted in which each term has a larger probable error than 1.

In calculating Table I., n was taken as 48. This gives 24 positive and 24 negative values in pairs, but I do not enter the signs on the dice, only the 24 values, leaving the signs to be afterwards determined by a throw of die III. It will be observed that the difference between the adjacent values in Table I. is small at first, and does not exceed 0.2 until the last three entries are reached. These, which are included in brackets, differ so widely as to require exceptional treatment. I therefore calculated Table II. on the principle of dividing that portion of the curve of distribution to which those entries apply, into 24 equal parts and entering the value of the ordinate at the middle of each of those parts in that table. Moreover, instead of entering the three bracketed values on die I. I leave blanks. Then whenever die I. is tossed and a blank is turned up, I know that I have to toss die II., and to enter the value shown by it.

The precise process I follow is to put 2 or 3 of dice I. into a small waste-paper basket, to toss and shake them, to take them out and arrange them on a table side by side in a row, squarely in front of me, but by the sense of touch alone. Then for the first time looking at them, to write down the values that front the eye. If, however, one of the blank spaces fronts me, I leave a blank space in the entries. Having obtained as many values as I want from die I., I fill up the blank spaces by the help of die II.

Lastly, the signs have to be added. Now as $24 = 16 + 8 = 2^4 + 2^3$, it follows that 16 of the edges of die III. may be inscribed with sequences of 4 signs in every possible combination, and the remaining 8 with sequences of 3 signs. Then when die III. is thrown, the several entries along its front edge, which are 4 or 3 in number as the case may be, are inserted in an equal number of successive lines, so as to stand before the values already obtained from the other dice.

The most effective equipment seems to be 3 of die I., 2 of die II., 1 of die III., making 6 dice in all.

Values for Die I.

0.03	···	0.51	···	1.04	···	1.78
0.11	···	0.59	···	1.14	···	1.95
0.19	···	0.67	···	1.25	···	2.15
0.27	···	0.76	···	1.37	···	(2.40)
0.35	···	0.85	···	1.50	···	(2.75)
0.43	···	0.94	···	1.63	···	(3.60)

Values for Die II.

2.29	···	2.51	···	2.77	···	3.25
2.32	···	2.55	···	2.83	···	3.36
2.35	···	2.59	···	2.90	···	3.49
2.39	···	2.64	···	2.98	···	3.65
2.43	···	2.68	···	3.06	···	4.00
2.47	···	2.72	···	3.15	···	4.55

Values for Die III.

+ + + +	+ − − +	− − + +	+ − +
+ + + −	+ − − −	− − + −	+ − −
+ + − +	− + + +	− − − +	− + +
+ + − −	− + + −	− − − −	− + −
+ − + +	− + − +	+ + +	− − +
+ − + −	− + − −	+ + −	− − −

CONCLUSION

All three of these works display ingenuity, and all were potentially useful as general techniques for investigating properties of statistical procedures, yet not one of them seems to have attracted a following. There are no doubt several reasons for this. Two of the techniques (De Forest's and Darwin's) made no claim to general usefulness, and their potential may have gone unnoticed. All three required specially made equipment, and while in no case should that have been an insurmountable difficulty, I believe that this should not be discounted—after all, simulation has only come into widespread use with the ease of implementation in modern computers. A contributing factor in this neglect may have been that those statisticians who might have tried Galton's device (say, Edgeworth, Pearson, Weldon) did not consider their problems as suitable for investigation in that way. The simulation techniques were all tied to the normal distribution, and all involved generating errors to be added to a signal. In some cases—least squares, for example—where that model might have been a reasonable supposition, Edgeworth and Pearson were usually capable of doing reasonably well with other approaches. The smoothing problems (really a sort of nonparametric regression) where the techniques appeared were of a quite different type. Nonetheless, one might expect that problems in meteorology or the estimation of growth curves would have provided potential applications for simulation. It is a distinct possibility, of course, that as historians of statistics take a closer look at the literature of the early decades of this century they will find that variants of these techniques were employed far more often than is now believed. In any event, these works of De Forest, Darwin, and Galton are eloquent testimony to the ingenious scientific energy of that era.

ACKNOWLEDGMENTS

I am grateful to R. L. Plackett and R. A. Thisted for comments, and to Agnes Herzberg for arranging for the photographs of Galton's dice, and to the Library of University College London for permission to have the dice photographed for this article. This manuscript was prepared using computer facilities supported in part by NSF Grants DMS-89-05292, DMS-87-03942 and DMS-86-01732 awarded to the Department of Statistics at The University of Chicago, and by The University of Chicago Block Fund. This research was supported in part by NSF Grant DMS-89-02667.

REFERENCES

Buffon, G. Le Clerc, Comte de (1735). [Solution de problèmes qui regardoient le jeu du franc carreau.] *Histoire de l'Académie Royale des Sciences de Paris* for 1733, 43-45.

Buffon, G. Le Clerc, Comte de (1777). Essai d'arithmétique morale. *Supplément à l'Histoire Naturelle* 4 46-148.

Chauvenet, W. (1891). *A Manual of Spherical and Practical Astronomy,* 5th ed. (2 vol.) Lippincott, Philadelphia.

Coolidge, J. L. (1925). *An Introduction to Mathematical Probability.* Oxford Univ. Press, Oxford.

Darwin, G. H. (1877). On fallible measures of variable quantities, and on the treatment of meteorological observations. *Philosophical Magazine* (Fifth Series) 4 1-14.

De Forest, E. L. (1873). On some methods of interpolation applicable to the graduation of irregular series, such as tables of mortality, etc. *Annual Report of the Board of Regents of the Smithsonian Institution for 1871,* 275-339.

De Forest, E. L. (1874). Additions to a memoir on methods of interpolation applicable to the graduation of irregular series. *Annual Report of the Board of Regents of the Smithsonian Institution for 1873,* 319-353.

De Forest, E. L. (1876). *Interpolation and Adjustment of Series.* Tuttle, Morehouse and Taylor, New Haven, Conn.

De Forest, E. L. (1877a, b). On adjustment formulas. *Analyst* 4 79-86; 107-113.

De Morgan, Augustus (1915). *A Budget of Paradoxes,* 2nd ed. (2 vol., David Eugene Smith, ed.) Open Court, Chicago. (1st ed., 1872.)

Duren, P. (1989). *A Century of Mathematics in America,* Part III. Amer. Math. Soc., Providence, R.I.

Galton, F. (1889). *Natural Inheritance.* Macmillan, London.

Galton, F. (1890). Dice for statistical experiments. *Nature* **42** 13-14.

Gosset, W. S. (1908a). The probable error of a mean. *Biometrika* **6** 1-24. (Reprinted in Pearson and Wishart, 1942, pages 11-34.)

Gosset, W. S. (1908b). Probable error of a correlation coefficient. *Biometrika* **6** 302-309. (Reprinted in Pearson and Wishart, 1942, 35-42.)

Gridgeman, N. T. (1960). Geometric probability and the number π. *Scripta Math.* **25** 183-195.

Hall, A. (1873). On an experimental determination of π. *Messenger of Mathematics* **2** 113-114.

Irwin, J. O. (1978). Gosset, William Sealy. In *International Encyclopedia of Statistics* (W. H. Kruskal and J. M. Tanur, eds.) 409-413. The Free Press, New York.

Jorland, G. (1987). The Saint Petersburg Paradox 1713-1937. In *The Probabilistic Revolution,* **1** (L. Krüger, L. S. Daston, M. Heidelberger, eds.) 157-190. MIT Press.

Kobbé, Gustav (1988). Behind the scenes of an opera-house. *Scribner's Magazine* **4** 435-454.

Kruskal, W. H. and Tanur, J. M. (1978). *International Encyclopedia of Statistics.* The Free Press, New York.

Lazzarini, M. (1901). Ricerche sopra una nuova espressione di π in funzione di soli numeri primi e sulla fattoriale di un numero. *Periodico di Matematico* **2** 49-68.

Lazzarini, M. (1902). Un'applicazione del calcolo della Probabilità. *Periodico di Matematico* **2** 140-143.

Muller, M. E. (1978). Random numbers. In *International Encyclopedia of Statistics* (W. H. Kruskal and J. M. Tanur, eds.) 839-847. The Free Press, New York.

Pearson, E. S. and Wishart, J. (eds.) (1942). *"Student's" Collected Papers.* Cambridge Univ. Press, London.

Perlman, M. and Wichura, M. J. (1975). Sharpening Buffon's needle. *Amer. Statist.* **29** 157-163.

Riedwyl, H. (1990). Rudolph Wolf's contribution to the Buffon needle problem (an early Monte Carlo experiment) and application of least squares. *Amer. Statist.* **44** 138-139.

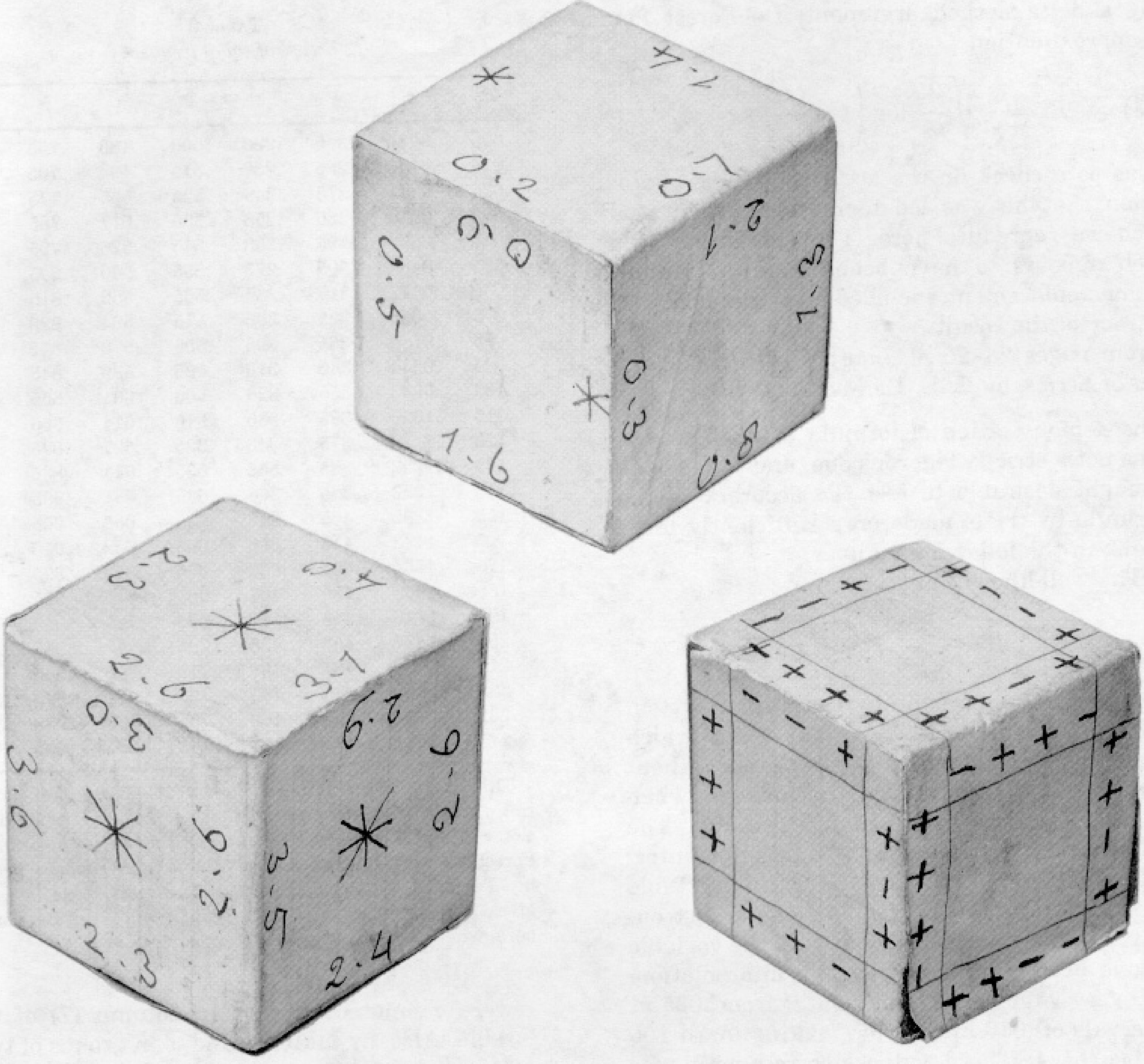
Stigler, S. M. (1978). Mathematical statistics in the early states. *Ann. Statist.* **6** 239-265. (Reprinted in Stigler, 1980, vol. 1, and in Duren, 1989, 537-564.)

Stigler, S. M. (ed.) (1980). *American Contributions to Mathematical Statistics in the Nineteenth Century.* Arno Press, New York.

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900.* Harvard Univ. Press, Cambridge, Mass.

Teichroew, D. (1965). A history of distribution sampling prior to the era of the computer and its relevance to simulation. *J. Amer. Statist. Assoc.* **60** 27-49.

Whitney, W. D. and Smith, B. E. (eds.) (1904). *The Century Dictionary.* Century, New York.


0.2
1.6
0.8
2.6
3.1
2.3
2.4
2.9
3.6