PROBABILITIES AND LIFE

Émile Borel

EXPLAINING MATHEMATICS

um, Banesh T. Hoffmann. $1.75

in Hubble. $1.75

.lloys, William Hume-Rothery. $2.25

ch, Desmond King-Hele. $2.00

cal Men, A. Albert Klaf. $2.00

al Men, A. Albert Klaf. $2.00

chnical Men, A. Albert Klaf. $2.00

unsky. $2.00

rinciples, Lewis Larmore. $1.65

The Science of Chance, Horace C.

Development of Their Scientific
..65

A Guide to Musical Acoustics, H. Lowery. $1.00

Visual Illusions: Their Causes, Characteristics and Applications, M. Luckiesh. $1.50

The Fourth Dimension Simply Explained, Henry P. Manning. $1.50

The Friendly Stars, Martha Evans Martin and Donald H. Menzel. $1.00

Matter and Motion, James Clerk Maxwell. $1.50

The Nature of Light and Colour in the Open Air, M. Minnaert. $2.00

The Concept of Heat and Its Workings Simply Explained, Morton C. Mott-Smith. $1.15

The Concept of Energy Simply Explained, Morton C. Mott-Smith. $1.25

Principles of Mechanics Simply Explained, Morton C. Mott-Smith. $1.35

Out of the Sky: An Introduction to Meteoritics, H. H. Nininger. $1.85

The Origin of Life, Alexandr Ivanovich Oparin. $1.85

Spinning Tops and Gyroscopic Motion, John Perry. $1.25

Catalysis and Catalysts, Marcel Prettre. $1.00

Modern Theories of Molecular Structure, Bernard Pullman. $1.00

Geometric Exercises in Paper Folding, T. Sundara Row. $1.00

An Elementary Survey of Celestial Mechanics, Yuri Ryabov. $1.25

Collected Papers on Acoustics, Wallace Sabine. $2.00

Atoms and Molecules Simply Explained: An Introduction to Chemical Phenomena and Their Applications, B. C. Saunders and R. E. D. Clark. $1.50

Microscopy for Chemists, Harold F. Schaeffer. $2.00

(continued on back flap)

PROBABILITIES AND LIFE

PROBABILITIES AND LIFE

By

ÉMILE BOREL

Translated from the French by

MAURICE BAUDIN

HEIDELBERG COLLEGE

DOVER PUBLICATIONS, INC.

NEW YORK

Originally published in French under the title *Les Probabilités et la Vie*, copyright © 1943 by Presses Universitaires de France.

Published in Canada by General Publishing Company, Ltd., 30 Lesmill Road, Don Mills, Toronto, Ontario.

Published in the United Kingdom by Constable and Company, Ltd., 10 Orange Street, London WC 2.

This Dover edition, first published in 1962, is a new English translation of the 4th edition of *Les Probabilités et la Vie*, published in the "Que Sais-Je?" Series by the Presses Universitaires de France. This English translation is published by special arrangement with Presses Universitaires de France.

Library of Congress Catalog Card Number: 62-5896

Manufactured in the United States of America
Dover Publications, Inc.
180 Varick Street
New York, N.Y. 10014

Contents

Introduction

The Single Law of Chance—Outline of the Work

1. The Single Law of Chance. Can there be laws of chance? The answer, it would seem, should be negative, since chance is in fact defined as the characteristic of the phenomena which follow no law, phenomena whose causes are too complex to permit prediction. However, following the lead of Pascal, Galileo, and many other eminent thinkers, mathematicians have built a science, the calculus of probabilities, the object of which is generally defined as the study of the laws of chance. Actually, the chief object of the calculus of probabilities is, as its name indicates, to evaluate the probabilities of complex phenomena in terms of the presumably known probabilities of simpler phenomena.

How can the calculus of probabilities permit the prediction of certain fortuitous eventualities? The mechanism of prediction is always the same and invariably brings in the *single law of chance*, of which we shall speak in greater detail, and which consists essentially of this: Phenomena with very small probabilities do not occur. It is then a matter of combining simple phenomena into complex phenomena whose probabilities calculated in terms of those of the simple phenomena are small enough for the application of the single law of chance.

We shall make use in this little book of certain results of the calculus of probabilities, but it is not at all necessary that the reader know the intricacies of the methods by which these results were obtained; he need only trust the mathematicians, just as an industrialist trusts his accounting department without feeling compelled to go over all the additions and multiplications.

The principles on which the calculus of probabilities is based are extremely simple and as intuitive as the reasonings which lead an accountant through his operations. The simple probabilities are sometimes deduced by arguments of symmetry: If a coin is flipped (as in the game of heads or tails), there is one chance out of two of turning up heads and one out of two of turning up tails. With a six-sided die, the probability for each side is *one-sixth*; there is one chance out of six of turning up *four*. Other probabilities, of a more complex nature, are derived from experiments or statistics. If out of 10,000 men of age 80, 1,087 die in the course of a year, we conclude that the probability that a man of 80 will die within a year is about 10.87 per cent, or 0.1087. Clearly, empirical probabilities are not rigorously established; all that is known are approximate values, which is also the case for all physical magnitudes that can be measured. However precise the measuring instruments, their accuracy is limited. The symmetry of a die or a coin is never perfect; and so the value 1/2 or 1/6 of the probability is only an approximation.

Once the simple probabilities are known, it is a matter of combining them. If two coins are flipped simultaneously or the same coin is flipped twice, the probability of turning up the same side twice will be equal to the product of *one-half* by *one-half*, or *one-fourth*. If two dice are thrown, the probability of obtaining the double six will be 1/6 multiplied by 1/6, or 1/36. But the probability of 6 and 5 with two dice will be 1/18, since one may have either 6 with the first die and 5 with the second, or 6 with the second and 5 with the first, and each eventuality has the probability 1/36.

By analogous procedures and observations the actuaries of insurance companies, using the mortality tables for men and women, are able to solve a problem such as the following: A couple, husband, age 60, wife, 55, pay 10,000 francs for an annuity to be paid till the death of the last survivor. What should be the amount of the annuity, for a given rate of interest, if no account is taken of the profit which the insurance company must retain to provide for general expenses and to set up reserves to guarantee that it shall meet its obligations in all eventualities? A more difficult problem, still solvable by the same principles, is how much the company must retain to be practically certain of fulfilling its engagements to its annuitants, even if some of them are lucky enough to live a very long time. To solve this second problem we must have recourse to *the single law of chance* which we have already mentioned.

This law is extremely simple and intuitively evident, though rationally undemonstrable: *Events with a sufficiently small* *probability never occur;* or at least, we must act, in all circumstances, as if they were *impossible.*

A classical example of such impossible events is that of *the miracle of the typing monkeys*,[1] which may be given the following form: A typist who knows no other language than French has been kept in solitary confinement with her machine and white paper; she amuses herself by typing haphazardly and, at the end of six months, she is found to have written, without a single error, the complete works of Shakespeare in their English text and the complete works of Goethe in their German text. Such is the sort of event which, though its *impossibility* may not be rationally demonstrable, is, however, so unlikely that no sensible person will hesitate to declare it *actually impossible.* If someone affirmed having observed such an event we would be sure that he is deceiving us or has himself been the victim of a fraud.

The case of the typist's reproducing the works of Shakespeare and Goethe without knowing them is so miraculous that no one can entertain any doubt about its impossibility. Less unlikely though still very improbable events could be imagined. Let us say that the typist wrote one verse of Shakespeare or Goethe, or only the first two words of one of their works. It is in such cases that the calculus of probabilities must intervene, since it permits the exact evaluation of the probability of the imagined event. We shall see later (Chapter Three) between what limits one is more or less authorized to regard this probability as negligible.

2. It is Repetition that Creates Improbability. If we analyze the case of the miraculous typist, we notice that the improbability comes from the fact that total success demands that a partial success be realized a very great number of successive times. The partial success consists of the fact that the first letter tapped by the typist is precisely the first letter of *Faust.* That success is not very probable since there are twenty-six letters in the alphabet, but it is not at all unlikely. The same may be said of the second letter which, by luck, might very well coincide with the second letter of *Faust*; the same of the third letter, and so on. Each of these partial successes, taken by itself, appears

[1] See Émile Borel, *Le Hasard* (Alcan), pp. 164–399.

quite possible; it is their almost indefinite repetition that creates improbability and rightly seems to us impossible.

One of the classical problems studied by the calculus of probabilities is precisely the probability of this or that result when the same test is repeated indefinitely. For instance, a coin is flipped and tails is considered the favorable result. What is the probability that the favorable event will be obtained 10,000 times out of 10,000 successive flippings? What is the probability of more than 6,000 favorable events out of 10,000 flippings? Calculation shows that these probabilities are so small that, according to *the single law of chance*, the respective events must be regarded as impossible.

3. Outline of the Work. We wish to study in this book applications of the calculus of probabilities to a number of questions chosen from among those which directly concern every man—questions which, for the most part, are related to either everyday living or to illness and death. We shall therefore leave aside the important applications of the calculus of probabilities to science, notably to the physical sciences.[2] Let us recall, however, that the importance of these applications, and the discoveries which they have produced, constitute one of the most solid proofs of the exactness of the results of the calculus of probabilities. Nor shall we develop the applications of the calculus of probabilities to the theory of games of chance, applications which were the origin of the calculus of probabilities and have remained one of its most attractive branches. We shall only refer to them from time to time, as we borrow from them simple examples to illustrate or make clearer certain results which we shall have to utilize.

The brief explanations which we have just given concerning the single law of chance and the miraculous typing feat suffice to suggest a preliminary difficulty to which our first two chapters are devoted. This difficulty is the following: The calculus of probabilities is an exact science, whose results are as reliable as those of arithmetic or algebra, *as long as it confines itself to the numerical evaluation of probabilities.* For thus one would obtain the probability of the miraculous feat of typing the works

[2] See my other works: *Le Hasard* (Alcan) and *Le Jeu, la Chance et les théories scientifiques modernes* (Gallimard).

of Shakespeare and Goethe. If these works comprise fifty volumes the size of this one, say about 10 million letters, the probability of the miraculous event which we have imagined is equal to unity divided by a number of more than 10 million figures. This result is as incontestable as that of any correctly done arithmetical operation. But in concluding from its extremely small probability that the typist's miraculous feat is impossible, by virtue of the single law of chance, we leave the domain of mathematical science, and it must be recognized that the assertion, which seems to us quite evident and incontestable, is not, strictly speaking, a mathematical truth. A strictly abstract mathematician could even claim that the experiment *need only* be repeated a *sufficient* number of times, namely, a number of times represented by a number of 20 million figures, to be sure, on the contrary, that the miracle will be produced several times in the course of these innumerable trials. But it is not humanly possible to imagine that the experiment can be so often repeated. If the dimensions of the universe are assumed to be equal to a billion billion light years, the number of atoms which it could contain, if it were full of matter, is expressed by a number of less than two hundred figures, and in the course of a billion billion years there are fewer seconds than a number of fifty figures would express. If, therefore, in that lapse of time, every atom of the universe were transformed into a typist and repeated the experiment every thousandth of a second, the total number of experiments would be much less than a number of three hundred figures. It is then clearly absurd to imagine experiments whose number would extend to more than a million figures; that is a purely abstract conception, a piece of mathematical juggling of no consequence, and we must trust our intuition and our common sense which permit us to assert the *absolute impossibility* of the typist's miracle which we have described.

Cases will present themselves, however, where intuitive evidence will be less clear and where it would nevertheless be legitimate, by virtue of the law of chance, to arrive at incontestable assertions. That these assertions do not possess the absolute truth of mathematical theorems must not be concealed; any such subterfuge would run the risk of justifying all the doubts as to their exactness. It must be well understood that the single law of chance carries with it a certainty of another nature than mathematical certainty, but that certainty is comparable to one which leads us to accept the existence of an historical character, or of a city situated at the antipodes, of Louis XIV or of Melbourne; it is comparable even to the certainty which we attribute to the existence of the external world.

This digression makes clear the nature of the preliminary difficulty to which the first two chapters are devoted. Simple common sense suffices to make anyone realize, more or less clearly, the peculiar character of assertions based on the calculus of probabilities; from there to expressing doubts about the exactness of these assertions is but a step. This step is readily made, as we shall see, by men whose psychology induces them to reject certain results deduced from the calculus of probabilities.

Our first chapter is concerned with the relations between the calculus of probabilities and the psychology of gamblers. The second chapter takes up the difficulties which arise in the minds of many quite reasonable men as soon as they consider the probabilities regarding human life.

In Chapter Three we try to specify the values of the probabilities which may and must be considered as practically negligible. We are thus led to define successively the negligible probabilities on the human scale, on the terrestrial scale, on the cosmic and the supercosmic scales. The chapter ends with some remarks on the definition of the probabilities of everyday life.

Chapter Four deals with the occurrence of events whose probability is very small, without, however, being absolutely negligible, when the number of tests becomes very large. We see that a very practical law, the law of Poisson, may be deduced from the single law of chance.

Chapter Five studies more intensively the probabilities of deaths already mentioned in Chapter Two, as well as the probabilities of illnesses and accidents. Finally, Chapter Six treats of some curious applications of the calculus of probabilities to certain problems concerning heredity in the human race.

We have relegated to appendixes certain developments which would have overburdened the text and which are not indispensable to the logical sequence of the ideas. Appendix One is devoted to a study of the repetitions of figures in six-figure numbers, numbers which interest all the clients of the national lottery. Appendix Two gives some arithmetical evaluations of Poisson's formula. Finally, Appendix Three contains statistical tables (tables of mortality, causes of deaths, accidents . . .) which will permit our readers to know precisely the value of the probabilities which directly concern their health and their lives.

One of my former pupils and the author of brilliant personal studies in the calculus of probabilities, Jean Ville, Professor in the University of Poitiers, has been kind enough to read carefully and correct the proofs of this book. I tender him my sincere thanks for his valuable collaboration.

CHAPTER ONE

Probabilities and General Opinion; Gamblers' Presumptions

1. Probabilities and Common Sense. There can be no doubt that some of the most reliable results of the calculus of probabilities seem to many people to be contrary to what is generally called common sense, that is to say, to popular opinion. I shall not undertake to analyze the somewhat obscure notion of common sense; I shall merely cite a brilliant page of Paul Valéry (*Regards sur le monde actuel*, Stock, 1931, p. 73):

> I do not feel comfortable when people speak to me of common sense. I believe that I have some; who would concede that he has none? Who could live another instant having found himself devoid of it? If I am confronted with it I become confused, I turn to the one who is in me, and does not have it and does not care, and who claims that common sense is the faculty which once enabled us to deny and brilliantly refute the supposed existence of the antipodes, a feat which it performs anew today when it seeks and finds in yesterday's history how to understand nothing of what will happen tomorrow.
>
> He adds that this common sense is an altogether limited intuition which derives from experiences neither precise nor planned, and which is mixed with a logic and analogies impure enough to be universal. Religion does not admit it in its dogmas. The sciences overwhelm it every day, upset it, obscure it.
>
> This critic of common sense adds that there is nothing to boast of in being *the most common thing in the world.*
>
> But I reply that nothing, however, can take away from common sense its great usefulness in disputes over vague notions, if there is no more powerful argument before the public than to invoke it for onself to proclaim that others are mad and that this precious gift, common though it may be, resides entirely in the speaker.

The lesson which seems to me derivable from these penetrating remarks is that when science shocks common sense it may be useful to find out why, and to seek arguments capable of convincing those who invoke common sense against science.

2. The Numbers of Lottery Tickets. Many people will refuse to buy a lottery ticket whose figures appear to them to possess, by their combination and choice, some exceptional property, as, for example, the number 272727 and, still worse, 222222. All who have thought about probabilities and about the methods used in drawing the winning tickets in the lottery know, however, that the probabilities of winning are the same for all the tickets, whatever their numbers. Yet many buyers of tickets will persist in saying, in the name of common sense: "Just the same, it is not possible for a number as singular as 222222 to win the first prize." The person who makes this assertion claims, furthermore, when the results are published, that in fact the first prize was won by a ticket bearing the number 825717 or 203409, and he concludes that his common sense did not deceive him and that he was right in not buying the number 222222, but the number 138615 which also, by the way, did not win.

There is no doubt that the probability that the number of the winning ticket will be made up of six identical figures is very small; it is equal to one one-hundred-thousandth, since there are 10 tickets in a million which bear six identical figures.[1] If there were 25 drawings a year, a ticket bearing six identical figures should win the first prize, on the average,[2] every 4,000 years. It is then quite probable that this remarkable event will not be observed by a man in the course of his life; but this in no way contradicts the calculus of probabilities, according to which the probability of winning is the same for all the tickets.

If any other number, or even a selection of ten numbers, is decided upon, it will be observed that generally none of these ten numbers wins. But if these are numbers without peculiarities it is not noticed at every drawing that these numbers did not win.

3. Numbers Formed by Means of Two Figures. The fact that the chances of all the tickets are equal will be better realized by studying a class of numbers quite peculiar, but also large enough so that the drawing of one of these numbers is observed now and then.

[1] We take into account the ticket whose number is 1000000, which is equivalent, if the drawing is made by means of six wheels, to the number 000000 formed by six zeros.

[2] See Chapter Four.

Such an example is furnished by the numbers composed of two figures, one of which may be zero. Such is, for instance, the number 233322, or 200200, or again, 55555, which must be written 055555. On the contrary, the number 55444 is formed with three figures, for it must be written 055444. The drawing is made by means of six spheres, each of which gives one of the six figures of the winning number.

It is easy to evaluate the number of tickets whose numbers are thus formed with two figures only.

If one of the figures appears 5 times and the other 1 time, there are $10 \times 9 \times 6 = 540$ numbers.[3]

If one of the figures appears 4 times and the other 2 times, there are $10 \times 9 \times \frac{6 \times 5}{1 \times 2} = 1{,}350$ numbers.

If, finally, each figure appears three times, there are $\frac{10 \times 9}{1 \times 2} \times \frac{6 \times 5 \times 4}{1 \times 2 \times 3} = 900$ numbers.[4]

There are then altogether $540 + 1350 + 900 = 2{,}790$ numbers in 1,000,000 which are formed with two figures only. If the ten numbers of one figure only are added to that, the total is 2,800, that is to say, about 1 in 357. The probability that such a number will win a given prize is then about 1/357. If we suppose the number of drawings and the number of prizes to be such that there are 360 important prizes a year (for instance, 30 drawings of 12 prizes, or 18 drawings of 20 prizes), the winning of an important prize by one of the peculiar numbers will be observed *on the average* about once a year.[5] It will therefore be a rare occurrence, but frequent enough, however, to be noticed by all those who look closely, after every drawing, at the list of the numbers which win the important prizes.

[3] The figure which appears 5 times may be any one of the 10 figures and the one which appears 1 time any one of the other 9 figures, which gives 90 possible choices; the figure which appears only once may occupy 6 different places; there are therefore in all $90 \times 6 = 540$ numbers.

[4] If each one of the figures appears 3 times, one of them may be chosen in 10 different ways and the next in 9 ways, but each combination, such as 3 and 4, is thus obtained 2 times (3 and 4, then 4 and 3); there are therefore 45 combinations, such as 4 and 3, and for each one of them 20 possible combinations: 444333, 443433, etc., that is, in all, $45 \times 20 = 900$.

[5] One may conclude, according to Poisson's formula (see Chapter Four), that out of 100 years, there will be about 36 when none of these numbers will win a prize, 36 when one of these numbers will win, 18 when 2 numbers will win, 6 when 3 numbers will win, 1 or 2 when 4 numbers or more will win.

In fact, if one took the trouble to consult many lists of drawings involving a million numbers (including the lotteries of the city of Paris, the *Crédit foncier*, etc.), one would easily observe that the proportion of winning numbers composed of only two figures quite conforms to the predictions of the calculus of probabilities.[6]

In Appendix One we shall study in more detail numbers of 6 figures from the point of view of the repetition of one figure in a number.

4. The Roulette Series. The problem of the series in a game such as roulette is extremely close to the one which we have just studied. It could even be regarded as identical if one used the binary (base two) system of notation. One may agree to represent the showing up of red by the number 0 and the coming up of black by the number 1 (we are speaking of a roulette having no zero). A sequence of spins of the wheel bringing up either red or black is then represented by a sequence of 0's and 1's, such as 10100100101110101. Such a sequence may be regarded as a number written in the binary system, and we may apply to these numbers the reasoning which we applied to the numbers written in the decimal system. We shall be led to admit that these numbers, diverse as they are, all have equal probabilities. A number composed exclusively of the figure 0 or 1 is very peculiar, and its turning up is very unlikely, especially if the number of figures is large, 30, for example; but the turning up of any other *specific* number of 30 figures is just as unlikely.

We disregard the binary system of notation and treat the question by direct reasoning, first discussing the delicate point where the results of the calculus of probabilities are contested in the name of common sense.

This delicate point is as follows: All players of roulette have observed that in a long series of spins reds and blacks are about the same in number. For instance, in 1,000 spins, one will observe 483 reds and 517 blacks, but never only 217 reds against 783 blacks. Most gamblers feel authorized to conclude from

[6] The proportion is very close, as could easily be seen if, as is the case for certain categories of bonds, their number is not exactly 1,000,000, but say, 500,000. If it exceeds 1,000,000, the figures for the millions may be disregarded.

this observation, exact in itself and besides conforming to the results of the calculus of probabilities, that if during a certain period they have observed more reds than blacks, the wheel has in a certain sense contracted a debt to black and will have to repay this debt by bringing up black oftener than red in the course of the next spins. In some cases the debt will even have to be paid off immediately. If a gambler, consulting the archives of roulette over a great number of years, observes that the longest run registered has been 24 reds, or 24 blacks, and that no run of more than 24 has ever been observed, and if some day he happens to observe a run of 24 reds, he will not hesitate to conclude that black must necessarily come up at the next spin, "since there has never been a run of 25."

To which Joseph Bertrand, together with all those who have probed the study of probabilities, replies: "The roulette wheel has neither conscience nor memory." It is paying it too great a compliment to imagine that it retains the memory of its errors and wants to make up for them.

It seems that "common sense" should be enough to persuade gamblers that successive spins of the wheel are independent of one another. It is indeed impossible to imagine a mechanism by which previous spins would modify the result of the next spin. But gamblers are influenced by an undeniable fact, confirmed by very numerous observations: in a large number of spins the reds are approximately as frequent as the blacks. They see no other way to explain this observed fact than to imagine the existence of an unknown mechanism playing the role of the conscience and memory of the roulette wheel and compelling the wheel to compensate for its errors.

An intensive study of the whole range of possibilities (quite analogous to the study developed in Appendix One for the decimal numbers of six figures) shows that if the combinations in which the reds are about as numerous as the blacks are more often observed than the combinations in which the reds are much more numerous than the blacks, it is solely because the former combinations are much more numerous than the latter, just as the six-figure numbers formed with 3, 4, 5 or 6 different figures are much more numerous than those formed of only one or two different figures.

It is not because the spheres of the lottery have a particular fondness for numbers made up of one or two pairs (one or two figures appearing twice in the number) that these numbers turn up oftener than the others, but because in a million numbers there are more than 680,000 containing one or two pairs. The same may be said of the distribution of reds and blacks in a series of spins of the roulette wheel (we disregard the zero). For instance, in a series of 30 spins the following results are obtained. The number of possible results of 30 spins is equal to the 30th power of 2, or a little more than a billion (exactly 1,073,741,824). In that billion possibilities, the various kinds of results turn up the number of times indicated below:

30 reds and 0 black ..	1 time
29 reds and 1 black ..	30 times
28 reds and 2 blacks ..	435 times
27 reds and 3 blacks ..	4,060 times
26 reds and 4 blacks ..	27,405 times
25 reds and 5 blacks ..	142,506 times
24 reds and 6 blacks ..	593,775 times
23 reds and 7 blacks ..	2,035,800 times
22 reds and 8 blacks ..	5,852,925 times
21 reds and 9 blacks ..	14,307,150 times
20 reds and 10 blacks ..	30,045,015 times
19 reds and 11 blacks ..	54,627,300 times
18 reds and 12 blacks ..	86,493,225 times
17 reds and 13 blacks ..	119,759,850 times
16 reds and 14 blacks ..	145,422,675 times
15 reds and 15 blacks ..	155,117,520 times
14 reds and 16 blacks ..	145,422,675 times
....	....
....	....
1 red and 29 blacks	30 times
0 red and 30 blacks	1 time

We have omitted the greater part of the second half of the table since it is obviously symmetrical to the first half. There is the same number of combinations with 17 reds and 13 blacks as with 17 blacks and 13 reds.

The run of 30 reds and the run of 30 blacks are unique combinations. Each is neither more nor less singular than any of the other particular combinations (the number of which surpasses a billion), but every particular combination is extremely unlikely. This would be the case of the combination consisting of alternate reds and blacks such that red would come up on every odd-numbered spin and black on the even-numbered spins.

Roulette players have never observed a run of thirty reds or thirty blacks and readily consider such a run impossible. In reality, if roulette could be played 1,000 times a day (1 whirl per minute for a little over 16 hours), a million days or about twenty-seven centuries, would be required to make a billion trials and thus have a good chance of obtaining a run of 30 reds (see Chapter Four, Poisson's law).

The case of at least 28 reds or at least 28 blacks are 2 (1 + 30 + 435) = 932, or less than one millionth of the total number of plays. Such an eventuality will be very rare, but still observable from time to time, if a patient player notes all the plays for a few years. At the rate of 1,000 plays a day, three years would suffice for more than a million plays.

The combinations involving at least 27 reds or at least 27 blacks number more than 8,000, or almost a hundred-thousandth of the total number of combinations. Such an eventuality will occur about one time in 100,000.

There are almost 63,000 combinations of at least 26 reds or blacks. This set of combinations will be realized about one time in 15,000 trials.

The combinations of at least 25 reds or blacks number almost 350,000. They will be realized on the average a little better than one time in 3,000 trials. The player noting 1,000 trials a day would observe them about 2 times a week.

If we proceed to at least 24 reds or blacks, the number of combinations slightly surpasses a million and the probability of observing one surpasses one one-thousandth.

As for the combinations of at least 22 reds or blacks, their number surpasses 10 million (about 17 million). The probability ranges between 1 and 2 one-hundredths.

Finally, the number of combinations of at least 20 reds or blacks is a little above 100 million, and the probability very close to *one tenth.* There are therefore nine chances out of 10 that, in a series of 30 plays, neither the number of reds nor the number of blacks will exceed 19. The average number being 15, it may be said, if the number observed is 19, that *the variation with respect to the average*, or more briefly, *the variation*, is equal to 4.

There are then 9 chances out of 10 that the variation be at most 4, that is, less than 5. It will be observed that 5 is the integral part of the square root of 30, the number of plays observed. This is a general law: *the probability of a variation equal to or greater than the square root of the number of plays is approximately equal to one tenth.*

5. The Law of Variations. It has been shown that the calculus of probabilities imposes no rigid laws to which chance should conform. Not only are relatively important variations possible, but to some extent they are probable and necessary. Anyone who observes with care and perseverance series of 30 plays will quite often note series containing more than 20 reds in 30 plays, and sometimes series containing more than 25 reds; but he will not observe any series containing 29 reds, and certainly none of 30 reds and no blacks.

If the number of plays of the series is much higher, for example 3,000 instead of 30, the probability of the variations remains the same, *provided the variations which are in the same proportion to the square root of the number of plays are paired,* that is to say, the variations which are 10 times greater for 3,000 plays than for 30 plays. Variations of 50 will therefore be quite probable, variations of 100 much less probable and variations of 150 practically impossible. If the number of plays were 300,000 it would be variations of 1,500 that would become extremely rare and, indeed, almost impossible. *The relative variation*, that is, the ratio of the variation to the number of plays decreases more and more as the number of plays increases. This is Bernoulli's law of large numbers, which is a simple arithmetical consequence of the single law of chance. The sets of 300,000 plays in which the variation is less than 1,500, that is to say, which contain fewer than 301,500 and more than 298,500 reds, are extraordinarily more numerous than the sets in which the variation is more considerable. The latter are not met with because, though very numerous, they are, compared to the others, extremely rare.

It is not only in games that one must keep in mind Joseph Bertrand's aphorism, "The roulette wheel has neither conscience nor memory." That is equally true of most common phenomena with which we are concerned in life, except in the cases where successive phenomena are not independent of one another. A well-known example of non-independent phenomena is that of rain and fair weather. A long series of rainy days increases the chances that it will rain again the next day, and a long series of fair days increases the chances of still another fair day. But if one observes rain and fair weather not for consecutive days but, for instance, at the same date every year, the rules of probability will apply. Meteorological statistics will tell us that in such or such a town, in the month of May, the number of rainy days is the same as that of fair days. There is then one chance in two that May 14 will be a rainy day. If we observe that date for a certain number of consecutive years, we shall be able to apply to these observations the results obtained for red and black at roulette. The fact that it has rained on May 14 five years in succession neither increases nor diminishes the chance that it will rain on that date the following year; they remain one in two.

If a telephone subscriber has noticed, after a careful observation, that from 2 to 6 in the afternoon his telephone is busy a total of two hours out of four, that is, half the time, I have one chance in two to find him free if I call. If I get a busy signal three times in a row, I still have one chance in two of reaching him if I call again. If I call every day, about once a month I shall get the busy signal five times in succession and more than once a year I shall get it eight times in succession. If we admit that a breakdown of the telephone automatically producing the busy signal would occur on the average once or twice a year, it is only after at least 8 consecutive busy signals that I can reasonably suspect that the telephone is out of order.[7] If I get 10 or 12 successive busy signals, the breakdown begins to be very probable. It will be almost certain if the busy signal is obtained 20 times in succession, at intervals of 5 or 10 minutes.

If I drive around in a town where many intersections are equipped with alternately red and green lights so that cars may use only one of the intersecting streets at a given time, I have one chance in two, at every crossing, to happen on a red or a green light. If my itinerary comprises twelve crossings, I must expect to meet, on the average, six red and six green lights. But if some day I have the bad luck to meet with

[7] The theory of the probability of causes is implicitly used here.

red lights at the first six crossings, I must not conclude that I shall have green lights at the other six crossings. If I make the same trip several times a day, I may well have, now and then, 10 or even 11, much more seldom 12, red lights or, on the contrary, 12 green lights. If one day I had the bad luck to be thus stopped at nearly every crossing, that will not in any way increase my chances of having a majority of green lights the next day. And yet, if I had the patience to compile the statistics for a whole year, I would find that the ratio of the total number of red lights to that of green lights is very close to unity.

CHAPTER TWO

Fallacies about Probabilities Concerning Life and Death

1. The Mysticism of Chance. One of the reasons why certain fallacies are so deeply rooted in gamblers is the very great importance they attach to winning or losing; gamblers are thus ever ready to receive favorably the most unreasonable suggestions if they believe they detect in them a way to conquer chance and insure victory. It is for the same reason that some singular superstitions on good and bad luck, the good and the evil eye, are quite common among theater people, actors or authors, whose success or reputation may depend on some incident in the course of a dress rehearsal. It seems to them that the merest trifle may bring about a brilliant triumph or, on the contrary, a failure which they will find difficult to overcome, and they are ready to make use of any means, even the seemingly most absurd, to turn luck in their favor.

But, important as may be winning or losing in gambling or success or failure in the theater, there is one possession to which men are still more attached—their own lives. That is why, in every question concerning more or less directly life and death, most men reason very badly, or rather cease to reason and let themselves be guided by their feelings or their prejudices.

The obscure and sometimes mystical ideas which many men entertain about chance and its role in life have been summed up with great genius by Rémy de Gourmont:

> Nothing is more expected than the unexpected, nothing, really, less surprising. What astonishes us above all is the logical course of facts. Man is ever waiting for a miracle, and is even chagrined if the miracle does not occur, or else he becomes discouraged. But the miracle often occurs. The humblest lives are but a sequence of miracles, or rather of chances. It will be argued that in reality there is no such thing as chance and that the word only proves our ignorance of the chain of causes. Since the chain of causes is undecipherable to our minds, we call chance all events whose occurrence we could not possibly discern. They are

formed, they happen, but we do not know and cannot know how or when. It is good that we cannot. Action is only possible if one is unconcerned and living is only an act of confidence in ourselves and the goodwill of chance.

We count on chance. There is no man, even among those most devoid of imagination, who does not grant it a place in his most obscure previsions. To rely on chance alone is madness, but to rely on it not at all is still worse. It is as unreasonable to despair as it is to hope. The impossible becomes possible at every moment of life. To be lost in a labyrinth a thousand feet under ground is a reason to hope, and one may with equal reasonableness lose all hope the day when, with a heart full of happiness, one looks at a life which is amiable and agreeable, attentive to our wishes.[1]

There are evidently many reasons, of which reason itself is not aware, that the applications of the calculus of probabilities to most problems concerning human life are often systematically ignored, sometimes even scorned and contested by the very persons they should interest.

There are, however, few results of the calculus of probabilities which have been better confirmed than those concerning mortality. For over a century life insurance companies have distributed to their stockholders dividends which are a tangible proof of the exactness of the computations of their actuaries, computations based on the calculus of probabilities and on mortality tables, that is to say, on statistical data.

This incontestable worth of data derived from carefully compiled statistics contrasts with the current prejudices against statistics. These prejudices come largely from what has been called "the individualistic mentality." Man is not pleased to be regarded as a simple unit, identical to other units; everyone is attached to his individuality and has a multitude of good reasons for considering himself really different from all others. Therefore, when the statisticians note a certain proportion of deaths among men of 40 years of age, every man of 40 will be ready to believe that this fact does not in any way concern him and that, being young and in good health, there is no reason at all for him to die within a year. Unless, of course, he rightly or wrongly considers himself gravely ill and about to die.

[1] Rémy de Gourmont, "Epilogues: 'L'Inattendu,'" *Mercure de France* April 15, 1906.

2. The Average Life. It is moreover perfectly true that the results of statistics must not be applied indiscriminately to all men of 40. Insurance companies demand a medical examination of all who apply for insurance. It would be more correct to distinguish in the coefficients of mortality derived from the tables the fraction which applies to the individuals regarded as in good health after a thorough medical examination and the fraction which applies to the individuals in whom the medical examination reveals certain diseases or hereditary taints (tuberculosis, cancer, syphilis, etc.). But, this distinction once made, there is no question that there is for every human being a certain probability that he will die within a year, a probability which depends on various elements, the most important being age and sex. This probability may be computed by means of the mortality tables of which we shall speak again.[2] These mortality tables permit the computation, at a given time and in a given country, of the average life of men and that of women, which is generally a little longer than the men's.

The *average life* of a given number of individuals is the arithmetical average of the length of life of every one of them. So defined, the average life can be calculated only when dealing with a group of individuals who have all died. Thus an investigation of the vital statistics of the nineteenth century would permit an evaluation of the average life of Frenchmen born in 1800 who died in France.

When dealing with a large population one may suppose, if the mortality tables are carefully compiled, that the ages at death of the mass of presently living persons will be distributed in proportions very close to those which would result from these mortality tables. That is why one may speak of the average life span of the inhabitants of a country at a given time.

One might imagine another method of evaluating the average life: take the arithmetical average of the ages at death of all the men and women who died in the course of a year. But one quickly realizes that this method would be accurate only if the population had remained practically stationary for a long period. If we note the number of deaths during the year 1941, the deceased of 20 years of age are persons who were born in 1921 whereas the deceased of 80 were born in 1861. If the

[2] See Chapter Five and Appendix Two.

population of the country considered had increased notably from 1861 to 1921, the number of deceased of age 20 would be too high relative to the number of deceased of age 80, so that the average life so evaluated would be below the actual life average.

Statistics show that in all civilized countries the span of the average life has notably increased during the last two centuries. This increase is due largely to the considerable decrease that the progress of hygiene has effected in the mortality rate of children of less than a year. It would be interesting, from different points of view, to consider the evaluated life average, not in regard to the number of births, but to the number of children who have attained one year of age. We shall return to this point in Chapter Five.

3. The Interpretation of the Mortality Tables. The few indications which we have just given suffice to show the importance which the data of the mortality tables and the evaluation of the average life have for us all, provided we fully understand them and do not exaggerate their significance. Every inhabitant of a country is undoubtedly concerned with the increase of the average life in that country, an increase which may result from the hygienic measures taken against the spreading of epidemic or contagious diseases, the building of hospitals and sanatoria, and so on.

I am well aware that an individualist could have recourse to his egotism and say: "I personally take all possible precautions to avoid contamination, I have an excellent doctor who watches over me and takes good care of me if I become sick. What do I care whether hospitals are built to which I shall never go, whether unhealthy quarters in which I am resolved never to live are cleaned up or not." Even from his purely selfish point of view this individualist is not right, for he cannot live isolated from the mass of other human beings, and he therefore risks becoming the victim of more or less direct contaminations which would have been avoided if certain diseases had been controlled by the progress of hygiene.

On the other hand, the excessive precautions taken against certain contaminations by people obsessed with their health cause at times unforeseen disasters. Cases are cited of persons who, having for many years drunk nothing but boiled water to avoid typhoid fever, die of that illness because of a single slip which would have been better withstood by one whose organism had been gradually accustomed to fighting microbes.

Joseph Bertrand devotes some very interesting pages[3] to the controversy which took place at the time of the discovery of the inoculation against smallpox, a quarrel in which some specialists in the calculus of probabilities took part. The problem which presented itself was the following: Vaccination killed one person in 100, but it reduced the probabilities, quite considerable at the time, of death by smallpox; must vaccination be advised or, on the contrary, forbidden?

Daniel Bernoulli, evaluating the life average under the two hypotheses (vaccination or no-vaccination), concluded that vaccination increased the life average by three years and declared that one ought not hesitate to practice it.

Joseph Bertrand, following d'Alembert's lead, has no trouble in showing that the evaluation of the average life is not sufficiently decisive and that other considerations must enter.

I have discussed elsewhere Joseph Bertrand's arguments.[4] The chief reason many people hesitate to let themselves be convinced by the evaluation of the life average is man's ignorance of the exact date when he will die, an ignorance which is one of the most important elements of his daily happiness. If the progress of science brought an end to that ignorance, and everyone could know when he would die, human mentality would be profoundly modified and everyone would attach a special importance to diverse circumstances which would change the date foreseen by the doctors for his death. But it is futile to reason on an unrealizable hypothesis; let us see things as they are.

It is quite apparent that, in facing the risks of diseases, men are divided into two categories, some passing from one category to the other, according to their humor, or belonging alternately to one or the other, according to the particular disease. One of the categories is composed of insouciants and the other of the obsessed. The former do not worry, and want to disregard the existence of microbes transmitting typhoid fever and the risks of contamination. They eat and drink as their parents

[3] *Calcul des probabilités* (Gauthier-Villars).
[4] *Le Hasard*, pp. 239 *et seq.*

and ancestors did, and believe that their robust constitution will protect them; if contamination happens, they accept it with fatalism. The obsessed, on the contrary, whose attention may have been aroused by something they have read or by a fatal disease observed among their friends, will think only, from morning to night, of precautions against the diseases which preoccupy them most (overlooking sometimes the risks of more dangerous or more frequent diseases). Neither category of men would care to know exactly the probabilities of contamination or death pertaining to a particular disease. These abstract figures would mean nothing to them. They are concerned only with the reaction of their personal sensibility to this or that disease. One will dread typhoid; another, cancer.

The statistical data which we reproduce in Appendix Three should make them all think and give them a fair appreciation of the real dangers to which we are all exposed.

CHAPTER THREE

Negligible Probabilities and the Probabilities of Everyday Life

1. Scientific Certainty and Practical Certainty. When we stated the single law of chance, "*events whose probability is sufficiently small never occur*," we did not conceal the lack of precision of the statement. There are cases where no doubt is possible; such is that of the complete works of Goethe being reproduced by a typist who does not know German and is typing at random. Between this somewhat extreme case and ones in which the probabilities are very small but nevertheless such that the occurrence of the corresponding event is not incredible, there are many intermediary cases. We shall attempt to determine as precisely as possible which values of probability must be regarded as negligible under certain circumstances.

It is evident that the requirements with respect to the degree of certainty imposed on the single law of chance will vary depending on whether we deal with scientific certainty or with the certainty which suffices in a given circumstance of everyday life.

If we are dealing with a scientific law, say Carnot's principle that heat cannot pass spontaneously from a hot substance into a cold substance, we have the right to demand that the probability of the phenomenon, impossible according to the law, be in fact extraordinarily small. If the law is to be called a physical law, the smallest infraction of it must not be possible under any circumstances, at any time, anywhere in the universe. We say briefly that the probability must be negligible on the supercosmic scale. Our preceding calculations of the number of atoms possible in a universe whose dimensions attain billions of light years and of the number of seconds in billions of centuries lead to an estimate of 10^{-500} (unity divided by a 500-figure number) as the negligible probability on the supercosmic scale. A probability so small may be taken to be 0 in the formulation of a scientific law. It goes without saying that the evaluation is a little arbitrary. Instead of the exponent 500 we could have written either 1,000 or only 200 or 300. In fact, the probabilities of a possible infraction of Carnot's principle in the kinetic theory of gases are very much smaller. They are equal to unity divided by numbers of millions of figures. Such probabilities must be regarded as *universally* negligible.

But in dealing only with human actions, the everyday life of any one of us, we shall see that a probability need not be so small to be rightfully disregarded in practice, i.e., considered as nil. We shall thus have to define negligible probabilities on the human scale, on the terrestrial and the cosmic scales. To these scales correspond degrees of practical certainty which do not attain the scientific and quasi-absolute certainty given by the supercosmic scale.

2. Probabilities which are Negligible on the Human Scale. We shall say that a probability is negligible on the human scale when it appears that the most prudent and reasonable men must act as if this probability was nil, that is, must run the risk of the occurrence of the event with which this probability is concerned, even if the event is considered by them a great misfortune. Such is the case, for example, of the death of the interested person or of a person particularly dear to him.

Let us take a simple example. According to peacetime statistics, the number of fatal accidents in traffic in a city like Paris, whose population counts several millions, is, on the average, one a day. For every Parisian who circulates for one day the probability of being killed in the course of the day in a traffic accident is about one one-millionth. If, in order to avoid this slight risk, a man renounced all external activity and cloistered himself in his house, or imposed such confinement on his wife or his son, he would be considered mad. The wisest persons do not hesitate to face every day a risk of death whose probability is one one-millionth. This is not, of course, a case in which the single law of chance warrants the assertion that the considered event will never occur; anyone who goes out every day into the streets of a great city knows very well that a fatal accident is *possible*. But he only thinks of it long enough to take, somewhat unconsciously, a few precautions to diminish his risk. He does not venture out to the middle of the street without first looking to see whether a car is coming; but he is not obsessed all day with the fear of a probable accident.

It is by comparing the number of accidents to that of the inhabitants of a great city that we came to proposing *one one-millionth* as a reasonable value for a negligible probability on the human scale. An analogous result would be obtained if we turned our attention to the number of times a man may perform in his lifetime gestures or very simple acts, such as tracing a letter of the alphabet with his pen, walking one step, or drawing one breath. The number of times is of the order of magnitude of a million in a few weeks, or months or years, according to the nature and frequency of the action. A prolific writer like Balzac wrote two or three million letters in a year. A professional typist could easily surpass that number. It may be concluded that the probability of writing one letter instead of another, whether it be done by a man using a pen or by a very expert typist, is certainly more than one one-millionth. If it were but one one-millionth, that is, one error in five hundred typed pages, everyone would agree to regard it as negligible and to declare that the typist had reached perfection.

3. Probabilities which are Negligible on the Terrestrial Scale. When our attention is turned not to an isolated man but to the mass of living men, the probabilities must be considerably smaller to be regarded as negligible. An accident which is quite improbable for one given man may be relatively quite frequent if all men are considered. Winning the grand prize in a lottery of a million tickets has a negligible probability for one holding a single ticket. If he is sensible, he will refrain from making plans for the future based on his winning the grand prize. This winning, on the other hand, is a certainty if all the tickets are sold and all the holders of tickets are considered: one of them must win.

If the number of human beings is estimated to be between one billion and two billion, a probability a billion times smaller than the negligible probability on the human scale must be regarded as negligible on the terrestrial scale, that is to say, a billionth of one one-millionth, or 10^{-15}, unity divided by a number of 15 figures. The same evaluation may be accepted if we consider all the men who lived in the historic past, that is, in the course of some hundreds of centuries, for their number is hardly a thousand times larger than the number of presently living men. Such probabilities may have to be considered, as we shall see in Chapter Seven, in handling certain problems pertaining to heredity in the human race.

The probability of obtaining red 50 consecutive times at roulette, or tails in the game of heads or tails, is 2^{-50}. If the approximate equality (very convenient in questions of this sort) $2^{10} = 10^3$ is used (in reality $2^{10} = 1{,}024$, a little more than 1,000), then 2^{-50} is equal to about 10^{-15}, that is to say, to the negligible probability on the terrestrial scale. In fact, if all presently living men were to spend all their time in playing a game such as roulette, at the rate of 1,000 games a day, or about 1,000,000 games every three years, only once every three years, on the average, would one of them obtain a run of 50 reds.

4. Probabilities which are Negligible on the Cosmic Scale. If we turn our attention, not to the terrestrial globe, but to the portion of the universe accessible to our astronomical and physical instruments, we are led to define the negligible probabilities on the cosmic scale. Some astronomical laws, such as Newton's law of universal gravitation and certain physical laws relative to the propagation of light waves, are verified by innumerable observations of all the visible celestial bodies. The probability that a new observation would contradict all these concordant observations is extremely small. We may be led to set at 10^{-50} the value of negligible probabilities on the cosmic scale. When the probability of an event is below this limit, the opposite event may be expected to occur with certainty, whatever the number of occasions presenting themselves in the entire universe. The number of observable stars is of the order of magnitude of a billion, or 10^9, and the number of observations which the inhabitants of the earth could make of these stars, even if all were observing, is certainly less than 10^{20}. A phenomenon with a probability of 10^{-50} will therefore never occur, or at least never be observed.

5. Probabilities which are Negligible on the Supercosmic Scale. Let us recall that the physical laws derived from statistical mechanics (and also the mathematical laws derived from the calculus of probabilities) have a still incomparably greater certainty which may be described by saying that the probability of the opposite event is negligible on the supercosmic scale. Such are the probabilities inferior to 10^{-n}, when n is a number of more than ten figures. If, for instance, there is in a liter container a mixture of equal volumes of oxygen and nitrogen, the probability that, at a given instant, all the molecules of oxygen be in the lower half of the container and all the molecules of nitrogen in the upper half is equal to 2^{-n}, n being the number of molecules.[1] It is negligible on the supercosmic scale.

An easy computation shows that if we evaluate the dimensions of our universe, that is to say, the distance of the farthest perceptible galactic nebulae, at ten billion light years, the volume of this universe is inferior to 10^{85} cubic centimeters and contains less than 10^{110} atoms, since the mean density is certainly inferior to 10^{25} atoms per cubic centimeter.

Let us imagine, with Boltzmann, a universe U_2 which contains as many universes U_1 analogous to ours as the latter contains atoms, then a universe U_3 which contains as many universes U_2 as U_1 contains atoms, then a universe U_4 which contains as many U_3 as U_1 contains atoms, and so on, repeating the same operation a million times, up to a universe U_N, with $N = 10^6$. This super-universe contains a number of atoms equal to 10 raised to the power 110 million, a number of one hundred and ten million figures. Let us imagine, on the other hand, a time T_2 containing as many billions of years as the billion years T_1 contains seconds, then a time T_3 containing as many times T_2 as T_1 contains seconds, and so on up to a time T_N whose index N is a million. Let us suppose that we do an experiment over as many times as there are atoms in the universe U_N and again as often as there are seconds in time T_N, that is to say, a number of times certainly inferior to 10 to the power 10^9. If the probability of the success of an isolated experiment is negligible on the supercosmic scale, an easy computation shows that the probability that the experiment will succeed a single time is so small that it can be disregarded. If we take as an example

[1] Since a gram molecule of gas contains $6{,}062 \times 10^{22}$ molecules, the number n of molecules contained in a liter is of the order of $3 \cdot 10^{22}$, and 2^{-n} is thus of the order of 10 to the power 10^{-22}.

the spontaneous separation of the oxygen and nitrogen contained in a liter, we can assert that this experiment will never succeed in time or space.

6. Probabilities and Everyday Life. Probabilities inferior to 10^{-6} or 10^{-15}, negligible on the human or terrestrial scale, are not often met with in practical life; but much larger probabilities must also be disregarded in the numerous cases where the event corresponding to such probabilities does not represent for us a grave misfortune, but merely a disagreeable incident. For instance, if it is a question of going out without an umbrella or a raincoat on a day when the weather is fair, the probability that it might rain could be evaluated from the statistics of the days when the weather was fair at 10 o'clock in the morning but it nevertheless rained in the afternoon. Without making the calculation, I believe I may assert that the probability is above one one-thousandth, at least in France. But unless a person is very frail, to the point that an unexpected rain might endanger his health and his life, we will not consider him careless if, on a fair day when nothing indicates a storm, he goes out without a raincoat or an umbrella.

It is unnecessary to accumulate examples; all men, even those who have never heard of the calculus of probabilities, evaluate probabilities without knowing it, even as M. Jourdain in *Le Bourgeois Gentilhomme* spoke in prose. Many of their decisions are influenced by the more or less confused notion they entertain of the probability of certain events. It might seem useless to know the calculus of probabilities, since mere common sense takes its place in most cases. I need no calculus to tell me to take my umbrella if rain is threatening and to leave it at home if the sun is shining. But it is equally true that, in some cases, I shall want to consult the barometer before I make up my mind, for the indications which it will give me will permit me to evaluate the probability that it will rain, with less chance of error, than if I merely glance at the sky through my window. I may also, if possible, consult a meteorological bulletin, or inquire about the direction and velocity of the wind. These supplementary precautions must not be neglected, if instead of merely risking being soaked, I go out to sea in a small sailboat, where bad weather may cause serious accidents.

Most men are as ignorant of the exact value of the probabilities which they use more or less consciously, as are young children and savage tribes of the exact value of money and of the prices of common articles. In one case as in the other the evaluation is dictated by subjective impressions, which often cause very gross errors. Before trusting a child regularly with money, it is well to teach him the worth of the articles which he may be led to buy. The same is true of probabilities. A man who has to run certain risks will gain by being precisely informed of their probabilities. Such is the case, for instance, of the probabilities concerning certain dangers or certain diseases. When someone has witnessed a serious accident, or observed among his friends some cases of contamination, he may be deeply impressed and unconsciously led to exaggerate the value of the probability that this accident or this contamination will occur again. If, on the contrary, the accident or the disease is one of which no example is known among acquaintances, he will tend to disregard the probability of its occurrence, though it may be relatively high.

The comparison which we made between ignorance of the value of probabilities and ignorance of the value of money and various merchandise may be pursued further. In very many cases one must run one risk or another, go out on foot or in a car, or else stay constantly at home at the risk of becoming anemic. A man with a delicate stomach must still eat, and must therefore choose among the possible inconveniences of the foods at his disposal.

The situation of one who does not know the probabilities is therefore similar to that of a man or child who has a limited sum of money and does not know the prices of merchandise. He runs the risk of spending all he has in a foolish manner. In the same way, ignorance of the probabilities may lead to running greater risks in trying to avoid lesser ones.

There is another analogy between prices and probabilities: the exact knowledge of prices is one of the elements in our decisions, but not the only one. If we have to choose between two articles of the same nature, one may appeal to us more than the other, and we may choose it, even if it costs more. It is nevertheless reasonable to inform ourselves of prices so that we may decide with full knowledge of what we are doing. If the price is ten times higher, we may hesitate to make so great a sacrifice to satisfy our fancy.

It is the same with probability. If we have serious reasons for wishing to travel quickly to another place, we shall consent to run greater risks of accident by traveling in a very rapid car or in an airplane. But if we knew that, under the circumstances, the risk of a fatal accident reaches one tenth, we would undoubtedly think twice before running this risk.

For the small child who does not yet know the value of money, the expressions ten francs, a hundred francs, a thousand francs are, if not equivalent, at least deprived of any precise significance. The same is true of the man who has never given thought to probabilities, when he is told of probabilities whose respective values are one tenth, one hundredth, one thousandth. A little thinking and practice is enough, however, to make one realize that there are many cases where it would be reasonable to run a risk with a probability of one thousandth, while it would be very unwise to run the same risk if its probability were one tenth.

Let us emphasize further the fact that, just as price is not the sole element of our decision when we make a purchase, probability alone must not dictate our decision in the matter of a risk. One of the reasons why certain minds mistrust the precision of mathematics is that they imagine that this precision may interfere with their freedom of choice. A sufficiently rich person may, of course, choose the articles he buys without regard to price, consulting only his taste. But when it is a matter of running a risk, especially if health or even life is involved, no one can afford to disregard certain probabilities, except when high considerations of morality and honor compel us to run a risk of death, great as it may be. In such cases it is best not to know the probability of the danger. But in everyday life the knowledge of the probability is a useful element in our decision, just as the knowledge of the price when we make a purchase, although this knowledge must not prevent us from taking other considerations into account before deciding.

7. Probabilities are Only Approximate. Probabilities must be regarded as analogous to the measurement of physical magnitudes; that is to say, they can never be known exactly, but only within a certain approximation. Moreover the degree of this approximation varies greatly with the nature of the probabilities. In cases where the probabilities may be evaluated by considerations of symmetry, the possible error in their evaluation is generally very small. Such is the probability of turning up a certain side of a die, or drawing a predetermined card from a new deck well shuffled and spread out on a table. A die is never a perfect cube, and the points marking the faces also introduce an asymmetry, but this asymmetry is very small, and the probability of each face turning up differs very little from 1/6. Similarly, the probability of drawing the king of diamonds out of a 52-card deck is very close to 1/52, although the cards cannot be rigorously identical, and differ furthermore in designs and colors. Errors in the evaluation of probabilities are necessarily more considerable when we deal with empirical probabilities derived from statistics. On the one hand, statistics are often imperfect, and are subject to systematic errors unavoidable and difficult to correct. We shall see examples of these apropos of the statistics of the causes of death. Further, statistics embrace only a limited number of cases, and the results obtained from them differ according as these statistics extend to a greater or lesser population, in a longer or shorter interval of time. Finally, probabilities vary generally with time, and one applies to the present year values of probabilities obtained from the statistics covering one or several of the preceding years.

Other probabilities are still more uncertain. They are those which are formulated from the impressions and recollections of even quite reliable persons. For instance, a doctor evaluates at nine out of ten the chances that a patient will recover from the disease with which he is afflicted, or a close follower of tennis tournaments evaluates at three out of four the chances that a certain player will win. We should be going too far, however, if, as some writers do, we denied any worth to evaluations of this nature, uncertain as they may be; but it is fitting that they be submitted to a serious scrutiny. The first point to ascertain is the sincerity of the person who evaluates the probability. We must ask ourselves whether there are not serious reasons to question his sincerity. A doctor may give to a patient's relatives an intentionally optimistic diagnosis. The habitué of the tennis courts may let his judgment be influenced by personal friendships or less respectable considerations in connection with wagers in which he has a personal interest. The surest way to determine the sincerity of a judgment of probability is to invite the person passing the judgment to make an important wager, provided, however, that he cannot exert any influence on the outcome of the contingent event on which the wager is made.

8. The Wagering Method. If a wager is made on the occurrence of an event with the probability p, the stakes must be equitably determined in the following manner. If Peter bets that the event will occur and Paul takes the opposite bet, Peter must put up a sum Ap and Paul a sum $A(1 - p)$. The total of the stakes, A, goes to the winner. If, for instance, Peter bets that he will throw a 6 with one die, he should put up 10 francs, and Paul 50 francs. The winner will collect the total of the stakes, 60 francs. If Peter throws a 6, he wins 50 francs; if he does not throw a 6, he loses 10 francs.

Let us now consider the case where the probability p is not, as in the case of dice, well known by the two bettors, but where Peter has declared that he estimates at p the value of the probability. If his estimate is too high, the sum Ap which he will have to put up will be too large, and the sum $A(1 - p)$ which his opponent will put up will be too small; the wager will be disadvantageous to Peter. If therefore Peter is suspected of exaggerating the value of the probability, as could be the case of an optimistic doctor who, wishing to reassure his clients, would exaggerate the probability of recovery, he could be made to reduce this exaggeration and return to a more exact evaluation by the obligation to wager a very large sum on the eventuality whose probability he has exaggerated. For instance, if a doctor declares that the chances of recovery are 9 out of 10 (a probability of 0.9), while they are really only 1 out of 2 (a probability of 0.5), and if he were to put up 90,000 francs in the hope of collecting 100,000 in the event of recovery, he would soon be bankrupt if this wager were often repeated. Out of 100 patients in the same situation we suppose that only about 50 recover. The doctor would put up altogether 9 million and collect only about 5 million if he bet on every one of his 100 patients.

The wagering method permits us to avoid the intentional errors which would be committed in the evaluation of probabilities, *when the nature of these errors is known.* But it is quite clear that, if instead of being optimistic, the doctor becomes pessimistic in some cases and evaluates the probability of recovery at 0.9, while it is really greater, equal, for instance, to 0.99, it will be to his advantage to accept a wager based on his evaluation. He will put up 90,000 francs to collect 100,000 in the event of recovery. If a single patient dies in 100 cases, he will have put up 9 million to collect 9,900,000 francs.

Is it possible by the wagering method to avoid the intentional errors committed by Peter in the evaluation of the probability, when these errors are not necessarily always in the same direction, that is to say, are sometimes an over-estimation, sometimes an under-estimation? It is possible, but on two conditions. First, that Paul has the right to dictate to Peter how he must bet. In the case of a sick person, Paul may bet as he chooses, either for the recovery or for the death of the patient. The second condition complements the first and is no less indispensable, namely, that Paul must be as competent as Peter in evaluating the probability, that he be himself a very good doctor if it concerns the treatment of a patient, and that he know in what direction Peter has altered the value of the probability, and place his bet accordingly. If Peter has exaggerated the probability of recovery, he will be compelled to bet on recovery. If, on the contrary, Peter exaggerates the probability of death, he will be compelled to bet on death. If Peter accepts these conditions, he will find himself by that very reason obligated to be absolutely sincere in his evaluation, since any systematic error would be ruinous for him.

It would moreover be natural enough for Peter, modest and prudent, to refuse to set a precise value on the probability of recovery, but merely affirm that in his judgment this probability is between 0.8 and 0.9, and that, under the circumstances, if he is forced to bet on recovery, he will demand that 0.8 be adopted, but if he is forced to bet on death, he will demand that 0.9 be adopted. Such an attitude would be perfectly correct, but Paul would be equally justified if he refused to bet under these conditions. It would mean that he agrees with Peter that the probability of recovery is between 0.8 and 0.9 and believes both wagers to be disadvantageous to him, whether Peter is to risk 80,000 francs against 20,000 by betting on recovery, or to risk only 10,000 francs against 90,000 by betting on death.

In fact, the method which we have just sketched to compel Peter to evaluate as correctly as he can certain probabilities has great analogies with the evaluation of probabilities of the rise or fall of the market value of a bond. The fluctuations are the results of offers and demands which represent the evaluation of a probability by the buyer and by the seller, each one estimating that this evaluation is advantageous to himself. Thus, the evaluation is a maximum for one of them and a minimum for the other.

9. The Combination of Wager and Auction. The method of auction sale often clarifies a buyer's exact evaluation of the worth of an object or a building offered for sale, for he stops bidding when the limit he has set himself has been reached. A similar method may be used, if Peter will accept it, to compel Peter to reveal precisely his evaluation of a certain probability. Let us return to the case where Peter is a consultant doctor who has been able to evaluate a patient's chances of recovery. We propose to find out whether he evaluates these chances at more than 50 per cent. We choose a contingent event with an exact 50 per cent probability, such as the game of heads or tails and we offer Peter an important gift or an intangible reward of considerable value to him and give him the choice between the following two eventualities: either he will receive the gift if the patient recovers, or he will receive the gift if a tossed coin turns up tails. He is clearly interested in choosing the eventuality with the higher probability in his opinion. He will choose the recovery of the patient if he considers the probability of this recovery to be above 50 per cent. If, on the contrary, he chooses the game of heads or tails, that will prove to us that he evaluates the probability of the recovery at less than 50 per cent. We may then repeat the test, using an eventuality with a 49 per cent probability. We may, for instance, with several decks of cards, whose backs are similar, make up a pack of 100 cards, 49 of which are red and 51 black. The probability of extracting a red card from this pack spread on the table after shuffling is 49 per cent or 0.49. If Peter prefers this probability to that of the case of recovery, it is because he evaluates the latter at less than 0.49. We may continue until Peter chooses the probability of the recovery when the other probability is only 0.43, whereas he had preferred the probability 0.44. We shall conclude that his true evaluation of the probability of the recovery is between 0.43 and 0.44. But, of course, true evaluation does not mean exact evaluation, since Peter is not infallible. Even if he is very skillful, it is quite doubtful that he can distinguish with certainty between probabilities as close together as 0.43 and 0.44. That is why it would be futile to seek a more exact decimal by diminishing the successive probabilities by a thousandth instead of a hundredth.

These evaluations of the probability of a particular event may be compared to the evaluation of a length or a weight which is made by a person with no means at his disposal for measuring them. If this person has a certain competence due to experience, his evaluation may be relatively exact, that is to say, contain two exact significant figures; perhaps three if the first figure is 1, as in the case of the height of a man evaluated in centimeters. Such evaluations do not have the value of a precise physical measurement done with good instruments, but are still preferable to absolute ignorance; it is the same with probabilities.

There is, however, a notable difference between these two sorts of evaluation, which comes from the fact that the methods which may be used to control these evaluations are different according as we are dealing with the evaluation of a measurable magnitude or with a probability. In the first case the control is easy: we need only measure with a good instrument and compare the result to the evaluation. Anyone may control his own evaluations and perfect himself in the art of evaluating at a glance the height of a man or a ceiling. In the case of a probability it is generally impossible to give a precise method similar to measuring a length by means of a meter rule. It is only by indirect and necessarily more complicated methods that the relative exactness of the evaluations made by this or that person of a certain category of probabilities can be determined.

10. The Control of the Accuracy of the Evaluations of Probability. It is not possible to control the accuracy of the evaluation of the probability of a single isolated event unless the evaluation is either extremely small or very close to 1, that is, amounts practically to an impossibility or to a certainty. But if we assert that an event has 9 chances out of 10 to occur or, on the contrary, 9 out of 10 chances not to occur, it may happen in either case that the event actually will or will not occur, and we cannot conclude that our evaluation was exact or not. An event may very well not occur, although its probability is 0.9, or occur, although its probability is only 0.1. Some writers believe they can solve the difficulty by refusing to examine it, that is, by denying the probability of an isolated event. I have discussed this point elsewhere[2] and shown why it does not seem to me to be acceptable. The notion of probability is a primitive notion, whose significance everyone grasps intuitively and which a scientific study can improve, as the geometrician improves the notions of straight line, plane, sphere, of which more or less crude examples are given by everyday experience.

Everyone knows perfectly well what he is saying when he asserts that a certain eventuality seems to him not very probable, rather probable, very probable, extremely probable, just as when he says that a person is small, average, rather tall and very tall. A certain experience permits us to substitute more precise numerical evaluations for these approximate evaluations and to say: I believe that this person is 1.60 m. tall, or I believe that the probability of this eventuality is slightly above one half, that is to say, this eventuality is more probable than the opposite eventuality.

The question is, how shall we know whether the evaluations of probabilities made by a given person are generally correct, and the evaluations of another grossly inaccurate. As the reader certainly suspects, it is the wagering method that will help us solve this problem, but, to avoid regrettable errors, the method must be used with circumspection.

We must observe that if a person makes an inexact evaluation of the probability, and we force him to wager on the basis of his evaluation, there are just as many chances that this wager will

[2] *Valeur pratique et Philosophie des probabilités* (*Traité du Calcul des Probabilités et de ses applications*, Vol. IV, fasc. III, Gauthier-Villars).

be advantageous to him or disadvantageous. That depends on the direction of the wager. If, knowing nothing about roulette, I assert that the probability of red is 3 out of 4 and that of black 1 out of 4, and another person as ignorant as I wagers 3 francs on red against my wager of 1 franc on black, this wager is advantageous to me and my error is profitable to me. Without going more deeply into the question, we may conclude that the wagering method, used without caution, would not single out the person who makes grossly inexact evaluations, since the cases where this inexactness will make the wagers advantageous to him will compensate the cases where the wagers will be disadvantageous.

That is no longer the case if one proposes to compare the respective skills of two persons separately evaluating the same probability who must act on their evaluations.

Let us suppose that Peter has evaluated at 0.5 and James at 0.7 the probability of the event which we call favorable (the recovery of a patient, the winning of a tennis match, the winning of a race by a designated horse). If they adopt for their wager the average value 0.6, James will be interested in betting for the favorable event and Peter in betting for the opposite event. For a total stake of 100 francs James puts up only 60 francs, while, according to his own evaluation, he should put up 70 francs, and Peter puts up only 40 francs, while, according to his own evaluation, he should put up 50 francs. If one of the two bettors, James or Peter, has made an exact evaluation of the probability, the wager is advantageous to him and disadvantageous to his opponent. But one may go further and note that, if both evaluations are inexact, the wager is advantageous to the bettor who has made the slighter error,[3] whether the errors are in the same direction or in opposite directions. For example, if the true value of the probability is 0.8, the wager of 60 francs against 40 is advantageous to James, while it is disadvantageous to him if the value of the probability is 0.4

[3] We evaluate the error by the *difference* between the true value and the value indicated by James or Peter. In this evaluation of the error we made use of the arithmetic mean. If we agreed—which may be preferable—to evaluate the error by the *relation* of the true value to the indicated value, we would have to choose as basis for the wager the geometic mean of 0.5 and 0.7, or about 0.59. The difference between the arithmetic and the geometric means is generally very small in practical cases: it will hardly ever happen that James evaluates the probability at 0.9 and Peter at 0.1.

or even 0.55 (here the errors have opposite signs).

If James and Peter make a single wager, it may, of course, be won by the one to whom it is, in principle, disadvantageous. But if they make a sufficiently larger number of similar wagers, the one who generally has the advantage[4] will win in the end. This is a consequence of Bernoulli's law of larger numbers. The probability that Peter will end by winning, when he makes a large number of disadvantageous wagers, becomes negligible as soon as the number of these wagers is sufficiently large.

The method of the wager applied to two persons shows which one is the more skillful in evaluating the probability. If a large number of persons taken in pairs are compared in the same manner, as for example, in the case of the diagnoses by many specialists of the same disease, it can be determined which one among them most correctly evaluates the probabilities, and it can be presumed that the evaluations of the winner of the tournament of wagers are as good as the present state of medical science permits.

[4] We say "generally" because, even if James is more skillful than Peter in evaluating, it may occasionally happen that an evaluation made by Peter will be better than the corresponding evaluation by James.

CHAPTER FOUR

Events of Small Probability. Poisson's Law

1. The Small but Non-Negligible Probabilities. It often happens that the probability of some events is not small enough to be negligible. One cannot then apply the single law of chance and assert that they never occur. But when the experiences are numerous, one may formulate certain approximate laws on the basis of the frequencies of these events. The probability of serious departure from these laws is sometimes small enough to permit the application of the single law of chance, and departures may then be regarded as highly improbable, or even practically impossible.

Let us consider a phenomenon with so small a probability that its occurrence may be regarded as exceptional. To be specific, we shall suppose the probability to be less than 1/30. If one trial is made every day, the phenomenon should occur, on the average, at most once a month. We shall suppose, on the other hand, that the probability is greater than 1/1000, although this hypothesis has no bearing on the results about to be cited, which remain true, however small the probability; but if this probability became too small, the trials we have in mind would be too numerous to be practically realizable.

2. Poisson's Law. Let us take, to be specific, a probability of 1/100, the probability, for instance, of winning the prize in a lottery of 100 tickets by the holder of one ticket. If the buyer of one ticket can repeat his experiment often, that is, frequently buy one ticket of a 100-ticket lottery, with the prize remaining the same, we have already asserted on several occasions, as an evident fact resulting from the very definition of probability, that the buyer in question, call him Peter, will win, on the average, one time out of 100. However, observation shows that if Peter repeats exactly 100 times an experiment consisting in buying one ticket of a 100-ticket lottery, he may well win one time and only one, but he may also not win a single time, and he may win two or more times. Poisson's theorem[1] reveals the probabilities of these various eventualities. According to this theorem, the probabilities that in 100 trials Peter will win 0 time, 1 time, 2 times, etc., are given by the following table:

Peter wins	0 time:	36.788%;	probability	0.36788
,,	1 ,,	36.788%	,,	0.36788
,,	2 times:	18.394%	,,	0.18394
,,	3 ,,	6.131%	,,	0.06131
,,	4 ,,	1.533%	,,	0.01533
,,	5 ,,	0.306%	,,	0.00306
,,	6 ,,	0.051%	,,	0.00051
,,	7 ,,	0.007%	,,	0.00007
,,	8 ,,	0.001%	,,	0.00001

It will be observed that the probability of winning one time and one time only is equal to the probability of winning 0 times. The probability of winning 2 times is 2 times smaller. That of winning 3 times is 3 times smaller than that of winning 2 times. That of winning 4 times is still 4 times smaller, and so on. The probability of winning 8 times is about 1 out of 100,000. That of winning 9 times would be 9 times smaller, that is, about one one-millionth and that of winning 10 times would be one ten-millionth. We are reaching the probabilities negligible on the human scale.

If 100 different persons make the same experiment as Peter, taking a lottery ticket 100 consecutive times, it can be asserted that among these 100 persons 36 or 37 will not win one single time out of the 100 drawings, about as many will win only once, while about 18 will win 2 times, 6 will win 3 times, 1 or 2 will win 4 times, and, rarely, 1 will win more than 4 times.

But, here too, the figures are only averages, and as always, variations from the average values are not only possible, but very probable, and must be regarded as the rule and not the exception, provided the variations are not too great.

3. The Variations. We have already said that the values of the variation which may be regarded as normal, that is, which are frequently seen, are those which are less than the square root of

[1] See in Appendix Two some mathematical developments relative to this theorem. While they are not actually indispensable to the understanding of what follows, they will surely interest those of our readers who have some mathematical knowledge.

the expected number. For example, out of 100 persons who have participated in 100 lotteries, it is expected that 36 or 37 will not win a single time (on the average 36.8). The square root of 36 is 6, it is therefore reasonable to expect that the number of persons who will not win a single time will be between 31 and 43. A variation double that of 6, which would correspond to less than 25 or more than 44 will be very rare, and a triple variation (less than 19 or more than 55) will be quite exceptional. The same results apply to the numbers of persons who would win one time and one time only.

As for the persons who would win twice, a variation of 4 from the average number 18 may very normally occur. The range would be between 14 and 22, it will seldom go down to 10 or up to 26. It must be considered quite exceptional for it to be less than 6 or more than 30.

Analogous results would apply to the cases of persons winning 3 times or more in a series of 100 lotteries.

These results show how disappointing is the gambler's profession, if one may so call the behavior of the person for whom gambling becomes a habit. The single prize of the lottery in which Peter continues to buy a ticket must be worth certainly less than 100 francs, if the ticket costs 1 franc, for the organizers of the lottery have to provide for certain expenses and even insure a profit. If the prize is worth 80 francs and Peter buys a single ticket 100 consecutive times, the probability that he will win one time is 0.37. In that case, his loss is 20 francs, since he bought 100 tickets at 1 franc apiece and won a prize of 80 francs. There is, on the other hand, a probability 0.37 that he will lose his 100 francs without winning anything. As for his chances of gain, they are as follows: about 18 chances out of 100 to gain 60 francs (2 prizes of 80 francs less 100 francs of tickets), 6 chances out of 100 to gain 140 francs, 1 or 2 chances out of 100 to gain 220 francs, the chances of a greater gain being extremely small. Analogous calculations would apply to the habitué of roulette who persists in playing a number other than zero (which he may change as he pleases, without altering the probabilities). Since roulette has a zero, he will win, on the average, one time out of 37, so that in 37 consecutive trials, the probabilities that he will never win or will win 1 time, 2 times, etc., are given by Poisson's table.

4. The Case where the Series of Trials is Repeated Several Consecutive Times. It is interesting to investigate what happens when the series of trials which we have imagined are repeated several times in succession: Peter buys 100 consecutive times a ticket in a 100-ticket lottery, or he plays 37 consecutive times a full number at roulette.

Let us suppose that Peter does not win a single time in the course of the first series. The probability of such an eventuality is 0.3679. If that occurs, the probability that Peter will not win in the course of the second series is not affected, and is also 0.3679. The probability that these two eventualities will occur consecutively, that is, that Peter will not win either in the course of the first series or in the course of the second series, is equal to the product of these two probabilities, or about 0.135. Such is the probability that in the course of these two series of 100 lotteries each, that is, in the course of 200 consecutive lotteries, Peter will not win a single time. If we consider a second series, also of 200 lotteries, the probability that Peter will not win is the same, that is, 0.135, and the probability that he will not win a single time in the course of the 400 consecutive lotteries (2 series of 200) is the product of 0.135 by 0.135, or about 0.018. This probability is close to 2 hundredths, and is not at all negligible.

The probability that Peter will not win in two series of 400, that is, in a series of 800, would be the square of 0.018, or about 0.0003, nearly 1 out of 3,000, a very small probability, but by no means negligible on the human scale.

One realizes here that the simple observation that Peter wins *on the average* one time out of 100 trials must be interpreted in the light of Poisson's calculations, if its significance is to be fully grasped. Peter should not interpret this formulation of an average value as implying the certainty that he will win the prize in 100, but in even several hundreds of consecutive trials.

It is the same when the probability in question is not the winning of a lottery prize, but an accident of which Peter runs the risk daily. For instance, Peter is a workman whose trade carries certain risks, an aviator, a locomotive engineer, or a truck driver. If the probability of an accident, according to the statistics of all the accidents which have happened to those exercising Peter's trade, is 1/1000 per work day, that means one accident about every three years (if one supposes that there are 333 work days in a year). But for every 100 persons exercising Peter's trade one must count about 37 who will have no accident in a first period of three years, and about 13 who will have no accident in two consecutive periods of 3 years. Such a proportion of exceptions must appear as quite natural, a simple consequence of Poisson's calculations of probabilities, and it is not necessary, in order to explain it, to bring in the differences between the probabilities among the various individuals.

The probability of such differences cannot, of course, be excluded a priori. A question is involved, which can only be answered by observation and experience. One may even regard as certain that these differences exist, since men are not all alike. Among truck drivers there are certainly some for whom the probability of an accident is less than the average, and others for whom this probability is, on the contrary, greater than the average.

The inequality of the probabilities associated with various individuals must obviously augment the proportion of those to whom, in a certain lapse of time, no accident has happened. We know that if the number of trials of each individual is equal to the denominator of the probability, that is, to 1,000 if the probability is 1/1000, one must expect that the event hoped for or dreaded will not happen to about 37 per cent of individuals. In the matter of accident, such will be the proportion of the spared individuals, aviators or truck drivers, for instance, who will have had no accident.[2]

A variation which does not surpass 6 per cent in either direction relative to the mean value must, of course, be regarded as normal, since it may be due to purely fortuitous causes. If the proportion of those who have had no accident is noticeably higher than 36 per cent, and reaches, for example, 45 or 50 per cent, it must be presumed that the variation is not fortuitous, but due to the fact that, among the individuals concerned, there are some for whom the probability is notably less, while for

[2] The probability 1/1000 is assumed to be calculated from certain statistics. It may refer either to a day's travel, a rather vague unit, since all days are not alike, or to a number of kilometers covered, say, a thousand kilometers. A thousand repetitions of the experiment will then amount to a million kilometers.

others, it may be more than the average. This last result would be confirmed, in accidents which are generally not fatal, as in automobile casualties, by the fact that the proportion of individuals who have had, in the period considered, more than 2 accidents, would be above 18 per cent and the proportion of those who would have had more than 3 accidents, would be above 6 per cent.

In the language of the calculus of probabilities, we sum up this increase of the proportion of cases where the number of accidents is 0, 2, 3 and the inevitably correlative decrease of cases where the number of accidents is equal to unity, which is the average, by saying that the observed *dispersion* is greater than the normal dispersion. It is a general law of the calculus of probabilities that, in this case, the material observed is not homogeneous: the probabilities are not equal for all individuals, but above the average for some and therefore below the average for others.

Can it happen that the observed dispersion is, on the contrary, less than the normal dispersion? It can when the observed phenomena are not independent of one another, as, for instance, in the case of contagious patients, or of observations bearing on numerous travelers using the same means of transportation. If a train loaded with travelers jumps the tracks, several hundred persons are simultaneously classified among those who have participated in an accident, and a fairly large number of them are sometimes killed or more or less seriously injured. A grave accident, assuming the proportions of a catastrophe, makes sometimes, in one day, a greater number of victims than the total yearly average. It is all the more so in maritime accidents.

However, in trains as in ships, there exists a certain independence between the chances of accident of two different persons. This comes from the fact that, except in rather particular cases (members of one family traveling together, inhabitants of the suburbs of a great city taking the same trains every day at regular hours), it is generally because of entirely fortuitous circumstances, which will not happen again, that travelers find themselves together in the same train or on the same boat. The probability that one of them will have another accident is independent of the analogous probability regarding one of his chance companions. It is not the same when one considers the probabilities of certain epidemic diseases, or of diseases whose frequency is influenced by very severe cold or heat. The probabilities vary then simultaneously for the inhabitants of a house, a neighborhood, a city or a region.

5. The Probabilities of Waiting. One of the practical problems which presents itself most frequently in everyday life is the probability of waiting, when the time depends on fortuitous circumstances, such as the number of clients at a window, or the more or less regular passage of a public bus.

Let us first consider the very simple case of a bus which passes at rigorously fixed intervals, say, every 20 minutes. If its schedule is not known, or no account is taken of it, the probabilities of arriving at a stop at any moment of the 20-minute interval between two successive buses must be regarded as equal. The average duration of the waiting will therefore be 10 minutes.

Let us now consider a slightly more complicated case. We suppose that the average interval between buses is still 20 minutes, but this interval is alternately 30 and 10 minutes. In other words, the hours of departure from the terminal are 12, 12:10, 12:40, 12:50, 1:20, 1:30, 2, 2:10, 2:40, etc. We continue to suppose that the traveler pays no attention to the schedule, either because he does not know it or because his watch is not right, or again, as is usually the case, because he has occupations or errands whose duration cannot be precisely evaluated, and he decides to take the bus as soon as he is free.

One might be tempted to reason as follows: when the interval between two buses is 30 minutes, the average waiting is 15 minutes, and when the interval is 10 minutes, the average waiting is 5; the average waiting, being alternately 15 minutes and 5 minutes, is, on the average, 10 minutes, the same as when the buses pass at regular 20-minute intervals. This reasoning is wrong, since it does not take into account a circumstance which a moment's reflection makes evident. The traveler who comes to a bus-stop at an arbitrary moment has many more chances to arrive during a 30-minute interval than during a 10-minute interval. Out of 4 times, he will arrive, on the average, 3 times during the 30-minute interval and 1 time only during the 10-minute interval. He will have therefore, out of four trials, an average wait of 15 minutes 3 times and an average wait of

5 minutes only once. The actual average waiting time will thus be:

$$\frac{1}{4}(3 \times 15 + 1 \times 5) = \frac{50}{4} = 12.5 \text{ min.}$$

that is, 12 and a half minutes; it is increased by the irregularity of the service.

One could attempt to solve an analogous problem, when the irregularities of the service are not systematic, but brought about by fortuitous circumstances, as frequently happens in the big cities where traffic is very congested. Although leaving the terminal at regular intervals, 10 minutes, say, the buses find themselves farther apart from each other by the time they reach the middle of their routes.[3]

It is somewhat difficult to submit the problem thus formulated to a rigorous calculation, since such a calculation could only be based on very precise hypotheses, on the probability of various irregularities (ahead of or behind schedule) which are regarded as possible. In the case of bus lines where departures are rather frequent, a fairly practical result will be obtained by admitting as an experimental fact that, when the average interval between buses is 10 minutes, intervals from 0 to 20 minutes are all about equally probable. The average waiting is 5 minutes when perfect regularity is achieved, and 10 minutes when irregularity is as great as it can be under our hypothesis (intervals whose lengths are alternately 10 and 20 minutes). An easy calculation leads to the conclusion that the average waiting is the arithmetic mean of 5 and 10 minutes, or 7 and a half minutes; it is increased 50 per cent by the irregularities of the service.

We have supposed, so far, that the traveler who waits always finds a seat in the first bus which comes along. To submit to calculations the cases where the coaches are full or can only accommodate some of the travelers, a number of hypotheses would be required which might be very arbitrary if they were not

[3] One may observe, in the case of buses, that a delayed bus will have to pick up a greater number of passengers at every stop, which will tend to increase its delay, whereas a bus ahead of schedule with respect to the preceding bus will pick up few passengers, so that its earliness will tend to increase. Thus it is that, on some lines, it frequently happens that a bus overtakes the preceding bus before they reach the terminal.

based on observation and statistics. The problem of waiting in the cases where the coaches are sometimes full or nearly full has some analogy with the problem of the waiting at windows, of which we are going to speak, restricting ourself to a very simple case, since the problem would be very complicated if one wished to study all the circumstances which can actually present themselves.

6. The Problem of Waiting at an Office Window. Let us first assume that the number of windows of a certain bureau, all identical, is strictly sufficient for the customers, who, we shall say to simplify matters, all require the same amount of time, 5 minutes. If the window is open 10 consecutive hours a day, it can serve 120 persons and 10 windows can serve 1,200 persons.

If the customers are less numerous at the beginning of the day, some windows will be idle part of the time, and the rush will be so great at the end of the day that the customers cannot all be served. If that situation occurs several times and becomes known, those customers who will consider themselves harmed by not being served by the end of the day, because of the rush, will make every effort to come at the beginning of the day. The consequence will be a greater-than-average rush in the morning and a more or less prolonged waiting. The problem is far from being a simple matter of probabilities, since it involves the psychology of the customers as well as numerous circumstances depending on the nature of the operations. It is only by greatly simplifying the hypotheses that we can submit it to calculations.

We shall suppose henceforth that there is only one window and that the daily clientele is below its capacity to the extent that, if the customers came at regular intervals, not only would there never be any waiting, but the windows would be idle during one fourth of the work day, that is, 2 hours (120 minutes in all) out of 8 business hours. It can, during the 6 hours of effective work, serve an average of 30 customers an hour, each staying two minutes, or 180 a day. But these 180 customers do not come at equal intervals. There generally are idle hours and rush hours. If, however, a large enough number of customers are free at all hours and are much afraid of having to wait, some of them will seek the hours known to be idle, and a certain equilibrium will be established. It is not absurd, therefore, to make the simple hypothesis that, for every one of the customers, all the hours of the work day are equally probable, everything happening as if everyone drew lots on the hour and minute of his coming. The problem may now be submitted to calculations, and, in spite of the simplicity of the hypothesis, the solution still is rather complicated. Appendix Two contains some precise calculations for readers interested in mathematics. We give here the results in the particular case just indicated.

Let us first make clear certain conventional phrases. The first customer to present himself at the opening of the window will be called the head of a series. If during the two minutes of his stay at the window, no other customer comes, the series is ended, and is composed of only one element. If, on the contrary, during the stay of the first customer, there come one or several customers, the series will only end when the window is again idle. It may comprise 2, 3, 4, etc. elements, each one consisting of one customer who makes use of the window for 2 minutes. If the series comprises 4 elements, its duration is 8 minutes. When it is ended, the first customer to come is again the head of a series, and so on, until the closing of the office. We shall suppose that closing time is held up a few minutes, if necessary, to accommodate waiting customers.

We have supposed that there are altogether 180 customers, whose business demands 6 hours, while the window is open 8 hours. The window remains unoccupied a quarter of the time it is open. Under these conditions the probability that a customer coming haphazardly in the course of the day will be the head of a series is precisely one fourth, hence the number of series will on the average, be equal to one-fourth of 180, or 45.

The respective probabilities that a series will comprise 1, 2, 3, or a larger number of elements can be calculated. These probabilities decrease rapidly at first and then much more slowly. By multiplying these probabilities by 45, the probable total number of series, we obtain the probable numbers of series of 1, 2, 3, 4, etc. elements. These numbers are 21 series of 1 element, 7 series of 2 elements, 3.5 series of 3 elements, 2.1 series of 4 elements, 1.4 series of 5 elements, 1 series of 6 elements. The number decreases very slowly after this point, since it is multiplied by about 0.9 every time the number of elements increases by one unit. It becomes 0.36 for 16 elements and 0.13 for 26 elements. But the sum of the probable numbers of series of 6 elements or more is equal to 10, which is far from negligible, and the probable number of the series of 29 elements or more is equal to unity. One may then expect a distribution like the following:

21	series of	1 element
7	,,	2 elements
3	,,	3 ,,
2	,,	4 ,,
2	,,	5 ,,
1	,,	6 or 7 elements
1	,,	8 elements
1	,,	9 ,,
1	,,	10 or 11 elements
1	,,	12 or 13 ,,
1	,,	14 to 16 ,,
1	,,	17 to 20 ,,
1	,,	21 to 25 ,,
1	,,	25 to 30 ,,
1	,,	31 to 40 ,,

But there may, of course, be variations relative to these average numbers. We sought only to indicate the general behavior of the phenomenon.

The number of series being 45 and the total number of elements 180, every series comprises, on the average, 4 elements. Let us recall that the total business time of the window, 8 hours, is equal to 4 times the number of idle hours; that is why the average number of elements is 4.

The average number which we have just calculated is the arithmetic mean of the number of elements of the various series, or, if one prefers, the average duration of these series (the unity of duration being 2 minutes). But in certain cases another definition of the average is preferable.

Let us consider a random customer. He will belong to a series, either as the head of it or as one of its other elements. When this series is terminated it will comprise a certain number of elements which may be called the number *observed* by the random customer. If we consider a large number of customers, each one will observe a certain number of elements in the series to which he belongs, and the average value of the series may be taken to be the arithmetic mean of the values observed by a large number of customers. The average so defined is evidently greater than the average which we previously calculated, since it is more probable that a customer taken at random will belong to a long series rather than to a short one. Calculations show that, in the present problem, the new average length is precisely the square of the previous one: 16 elements instead of 4. If a customer comes in at random and belongs to a series of 16 elements, he has equal chances to occupy any of the places between the first and the sixteenth. The number of those preceding him is between 0 and 15, an average of 7.5. This is the most precise and the most general answer to the problem of the waiting at the window. A new calculation would be required to determine the average duration of the waiting.

If the window were idle for a length of time equal to half of the business hours (not to a fourth as was supposed) the average length of the series would be 2, according to the first method of calculating, and 4, according to the second method. Every customer would have an average of 1.5 predecessors in the series to which he belonged. Half of the customers would be heads of series and have no predecessor.

If, on the contrary, the window were idle for only one tenth of the business hours, the average length of the series would be ten, according to the first method of calculating, and 100, according to the second method. There could be, if not every day, at least quite often, series above 100. The average daily number of series being only 18, observations bearing on several days would be required to verify our results concerning only averages.

CHAPTER FIVE

Probabilities of Deaths, Diseases and Accidents

1. Probabilities of Deaths. As early as the eighteenth century statistics of deaths in relation to age began to be seriously established.[1] During the nineteenth century these statistics became quite accurate in all civilized countries. Moreover, the life insurance companies, whose number and importance did not cease to grow, established very precise statistics on their clientele. In these statistics the companies differentiate between two categories of clients, according to the nature of the insurance policy. In certain policies the death of the insured is an advantageous event, if not for him, at least for his heirs, and therefore disadvantageous to the company which must pay a considerable sum to them. In other contracts, on the contrary, it is the prolonged life of the insured which is advantageous to him and disadvantageous to the company which must pay him a life annuity. In the language of insurance companies, the first category are the *insurants* and the second the *annuitants*. The mortality of the annuitants is obviously lower than that of the insurants, even though the companies demand a medical examination of the latter and refuse to insure them if the examination is not favorable. A man who knows he is sick or frail does not easily decide to put himself in the category of annuitants by sacrificing a considerable capital in exchange for a life pension.

The Mortality Tables of French policy holders, reproduced by the *Annuaire du Bureau des Longitudes* from the publications of the insurance companies, are given in Appendix Three. The Appendix also contains other tables of the general statistics of France and tables of survival by generations, accompanied by comments on the differences between these various tables. We remark here only that the tables of insurance companies refer to a somewhat selected population, since the insured have passed a medical examination before a company's doctor, and it is to the interest of the annuitants to consult their own doctors before

[1] Deparcieux' Table, 1746.

taking a policy. But these examinations take place only once, when the policy is taken, and the duration of the contract is often very long. In the course of it insurants or annuitants may have some grave illnesses which considerably increase their probabilities of dying within a year, compared with the mean probabilities of the group of men or women of the same age.

Let us insist a little on the distinction which must be made between the mean probability of death in the course of a year for a man of forty and the analogous probability when it is known that the man of forty is presently well and runs no exceptional risks in the exercise of his profession or his mode of living.

2. The Meaning of the Mean Probability. Let us consider, to fix our ideas, the men of 40, according to the Table of Survival of French insurants (French companies, 1895). Out of 711,324, 6,938 die in the course of a year, a little less than 10 out of a thousand. If we adopt the figure 10 out of 1,000, the mean probability of mortality in the course of the year is 0.01, or one hundredth, for a man of 40 *on whom there is no other information available*, and who may legitimately be considered as chosen at random among the men of 40. Take, for instance, the totality of the men who reach their 40th year in January: if they number 30,000, it must be regarded as probable that 300 will die before the age of 41. The difference which may be observed between the actual number of deaths and the average number of 300 calculated according to the probability will be relatively small, of the order of magnitude of the variations which occur when a simple experiment is repeated a great number of times, such as the throw of a die or the drawing of a number out of an urn.

If account is taken of the possibility of exceptional events which increase the general mortality (war, epidemic, abnormally cold winter), the variations concerning mortality are sometimes greater than those of simple fortuitous events.

It may, then, be assumed, in the examples we have picked, that the number of deaths will be above 250 and below 400. But this supposes that the experiment embraces 30,000 persons of 40 really taken at random. Such is the case if, instead of those born in January, we choose those whose names begin with A or B. But if one chooses 30,000 office workers between 40 and 41, working on January 1, one must expect a definitely lower mortality, since their being active proves that they do not on that date suffer from a grave illness, such as tuberculosis. Moreover, one may expect that the probabilities of certain illnesses or causes of accident are less for office employees than for workmen or farmers.

The variations of the probability in dealing with a restricted category of persons, would be still greater if, instead of the probability of death *in the course of a year*, we spoke of the probabilities of death *in the course of a day*, within 24 hours, from noon today to noon tomorrow.

For the aggregate of men of 40 this probability is 365 times less than for the year, that is, one may expect an average of 10 deaths out of 365,000 persons instead of 10 out of 1,000. In a large country, if the number of persons of 40 were 730,000, the average daily mortality among persons of that age would be 20. But this number would evidently be very much smaller, if we considered only the persons of 40 who, today at noon, are in good health, and will not, in the next 24 hours, run any exceptional risk of accident (a long trip by plane or car, a dangerous acrobatic exhibition, etc.). There are very few illnesses that kill in 24 hours without previous warning, and many fatal accidents grant to their victims a few hours or days of survival. It would therefore be a gross exaggeration to evaluate at 1 out of 36,500 the probability of death in the course of 24 hours for a person in good health and not running any exceptional risk. It may be asserted that the probability is very much lower, although a precise evaluation, or even a definition of it, is somewhat difficult. What is meant by "a person in good health"? Must the person's assertion suffice, or is a medical examination required? On the other hand, what are the risks of accident that must be considered either normal or exceptional?

It might be interesting, however, to distinguish better than has been done so far the probabilities of global survival at a certain age for the whole of a population from the probabilities relative to persons of this age who are in good health and run no exceptional risks. The statistics of deaths classified according to their causes would be one of the most important elements of this study.

3. Deaths According to their Causes. The enforcement of the laws requiring declaration of the causes of death has made great progress in France since the publication of the first edition of this book [1943], largely because of the development of social insurance; as a consequence the doctor is called in almost every grave illness. Whereas in 1936, out of 642,000 deaths, there were about 131,000, more than 20 per cent, due to unspecified or ill-defined causes, in 1948 the number was only about 35,000 out of 506,000 deaths, less than 7 per cent. We have therefore replaced the tables which we had given for 1936 by the statistics published for 1948, which is the year when the number of deaths was the lowest for the half-century period 1900–1949.

We have reproduced the new classification of the causes of death. Like all classifications, this one cannot be perfect, and it must be conceded that in many cases a doctor may legitimately be puzzled. Here is a patient suffering from a sort of tuberculosis which may be presumed curable. However, after exposure to excessive cold, he dies of bronchitis or pneumonia. Is the death attributable to tuberculosis or to the accidental illness? A similar question often presents itself with syphilitic patients. According to the specialists, the number of deaths whose real cause is syphilis is in reality much higher than the statistics indicate, for they name more readily an accidental cause in numerous cases where the latter would probably not have brought about death, had not the victim been syphilitic.

We have given, regarding the most frequent causes of death, the distribution of deaths in relation to ages as it appears in the *Bulletin de la statistique générale de la France.* It was, of course, useless to give the distribution for cause XV (illnesses peculiar to the first year of life). As for cause XVI (senility, old age), let us state here that, for the men, there are 488 deaths between the ages 50 and 69, 4,722 between 70 and 79, and 9,578 between 80 and 99, and for the women 572 between the ages of 50 and 69, 6,188 between 70 and 79, and 16,954 above 80. These figures are explained by the greater longevity of women.

We lacked room for the distribution by provinces of deaths from various causes. This breakdown is rather instructive. It emphasizes important differences sometimes due to the variety of climates or the existence in some provinces of specialized hospitals, but more often to differences in terminology among the doctors of various regions. The proportion of deaths whose causes are not declared or ill-defined varies much according to regions.

CHAPTER SIX

Application of Probabilities to Certain Problems of Heredity

1. Heredity and the Chromosomes.* According to theories generally accepted by biologists and confirmed by numerous experiments, the phenomena of heredity are linked to the existence, in every individual, of a number of pairs of chromosomes (24 in the human species). These pairs are differentiated from one another and may therefore be numbered. In every child the chromosomes of a given pair, say, the seventeenth, are formed from one of the chromosomes of the father's seventeenth pair and one of the mother's seventeenth pair. It is as if the child drew lots and had one chance out of two to choose either one of the father's two chromosomes and either one of the mother's two chromosomes, for the seventeenth and for every one of the other pairs. The number of choices among 48 pairs is 2^{48}, more than 250,000 billions. When two brothers or sisters are not twins from the same egg, in which case they have the same chromosomes, and are exactly alike, the probability that the choices will all be the same is extremely small, the quotient of unity over 250,000 billions. It is not probable that such an event happened on earth since the beginning of the human species.

Although the precise role of the chromosomes in the determination of physical, intellectual and moral characteristics of every individual is still little known, it seems that the presence in two individuals of certain groups of identical chromosomes is enough to create very striking resemblances or analogies between them. A single chromosome may sometimes determine a characteristic important enough to attract immediate notice. Such is the case of certain hereditary taints. This shows the significance of the study we shall presently make of the probabilities relative to the simultaneous presence of a chromosome in individuals who have one or several ancestors in common, *resemblances between brothers, uncles and nephews, first cousins*, etc.

*See note, page 63.

2. Chromosomes Common to Brothers and to Cousins. Consider first two brothers having the same father and mother. All the chromosomes of one, whom we call *A*, come from either the father or the mother, that is, from one of the two parents common to *A* and his brother *B*. If we pick out a given chromosome of *A*, there is one chance in two that it will be found in *B*. Out of the 48 chromosomes of *A* an average of 24 will be found in *B*.

We have implicitly supposed that the father and mother of the two brothers are not related, that is, have no common chromosomes.

Consider now an uncle and a nephew. The nephew's father is understood to be a full brother of the uncle, that is, they have the same father and mother, but the nephew's mother is not related to her husband. Under these conditions, the uncle and nephew have two common ancestors, the uncle's parents and the nephew's paternal grandparents. The uncle's chromosomes come from one of these two common ancestors, but there is only one chance in four that any one of these chromosomes will be found in the nephew, since the latter is separated from the two common ancestors by two generations (including his own).

The uncle and nephew have in common an average of 12 chromosomes.

If we consider first cousins, we shall specify that their fathers are full brothers and their mothers unrelated to each other or to their husbands. Under these conditions, they have two common ancestors, their two paternal grandparents. A chromosome of one of the cousins has one chance out of two of coming from the common ancestors, and there is one chance out of four that it will be found in the other cousin. The probability that a chromosome of a given rank will be common to the two cousins is therefore $1/2 \times 1/4 = 1/8$. Out of 48 chromosomes they have an average of 6 in common.

We take now the case of first cousins whose fathers are brothers and whose mothers are sisters. They have four common ancestors, any chromosome of one of them comes from one of these four ancestors. But every one of these ancestors is separated from his grandson by two generations, that is, two choices. One of their chromosomes has only one chance out of

4 to be found in the grandson. The two cousins have in common an average of 12 chromosomes, an average of 3 of these chromosomes coming from each of the four common ancestors. The difference between the case of these two cousins and that of two brothers who have also 4 common grandparents is explained by the fact that, in the case of the two brothers, the *parents have already made a choice among the chromosomes of* the grandparents, and this choice is the same for the two brothers.

If a characteristic is linked to a single chromosome, this characteristic is common to two brothers one time out of two, to uncle and nephew one time out of 4, one time out of 8 to two first cousins (one time out of 4 to double first cousins).

If a characteristic is linked to the simultaneous presence of two chromosomes, it will be present only one time out of 4 in two brothers, one time out of 16 in uncle and nephew, one time out of 64 in first cousins.

It would seem then that a few precise statistics would suffice to determine, a posteriori, from resemblances between brothers, etc., what characteristics depend on one, two or several chromosomes.

3. A Few Words on a More General Case. We have supposed, as is most frequently the case, that two brothers have the same two parents. It would be easy to treat the more general case in which the common ancestors are not necessarily at once a father and mother. Let us take, for instance, two cousins who have in common a grandfather and a great-grandmother, the other common ancestors being exclusively the ancestors of these two.[1] Given a chromosome of one of the cousins, there is one chance out of 4 that it comes from his grandfather and one chance out of 8 that it comes from his great-grandmother, these two eventualities are mutually exclusive. In the first case, there is one chance out of 4 that the chromosome exists also in the second cousin, and in the second case, there is one chance

[1] Given Paul and James, the two cousins. Paul is the son of Peter and Jean, and James the son of Henry and Bertha. Peter and Henry are the sons of the same father and not of the same mother. Jean is the daughter of Edward, and Bertha the daughter of Marguerite. Edward and Marguerite have the same mother and not the same father.

out of 8. The probability that the chromosome will be common to the two cousins is

$$\frac{1}{4} \times \frac{1}{4} + \frac{1}{8} \times \frac{1}{8} = \frac{5}{64}.$$

An analogous formula would apply, whatever the number of common ancestors, who may well not correspond to the same generation for both cousins. Generally, the ancestor of order a_1 of A is supposed to be the ancestor of order b_1 of B (if $a_1 = 1$, we are speaking of the father, if $a_1 = 2$, of the grandfather, if $a_1 = 3$, of the great-grandfather, etc.), the ancestor of order a_2 of A is the ancestor of order b_2 of B, the ancestor of order a_3 of A is the ancestor of order b_3 of B, etc. The probability that a chromosome is common to A and B is

$$\mathrm{P} = \frac{1}{2^{a_1+b_1}} + \frac{1}{2^{a_2+b_2}} + \ldots + \frac{1}{2^{a_k+b_k}}.$$

If we speak of two double first-cousins, that is, having four grandparents in common:

$$a_1 = b_1 = 2; \; a_2 = b_2 = 2; \; a_3 = b_3 = 2; \; a_4 = b_4 = 2,$$

and, from the formula,

$$\mathrm{P} = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}.$$

There remains the case where one of the common ancestors must be considered a multiple ancestor by one of the descendants (when cousins are married to one another). Without going into details, let us indicate that every individual has two ancestors at the first generation (parents), four at the second generation (grandparents), eight at the third generation (great-grandparents), etc. If among the 16 ancestors of the fourth generation, one person figures two times, he must be counted twice, that is, must be given two numbers (equal to each other) a_1 and a_2 in the above calculation. If among the 32 ancestors of the fifth generation, one person figures 3 times among the ancestors of A and 2 times among the ancestors of B, there will be 3 numbers a each equal to 5 and 2 numbers b each equal to 5, which gives $3 \times 2 = 6$ sums $a + b$ equal to 10, that is, 6 terms each equal to $\frac{1}{2^{10}}$.

We leave to the reader the study of the more complicated cases which may present themselves. The one in which one ancestor figures two or several times in the ancestry of one individual, with possibly different ranks, offers no particular difficulty. A less simple case is that of our two individuals A and B having common ancestors who are related to one another, that is, who themselves have common ancestors. The simplest example of this is the case of two brothers whose father and mother are more or less distant cousins. Still more complex cases, in which, to be thorough, one would have to go back an almost indefinite number of generations, are frequent in isolated villages where for centuries a small number of families have intermarried, with very little new blood from outside.

4. Application of the Single Law of Chance. All the results which we have just indicated concerning heredity can be translated into the language of probabilities. They cannot therefore lead to any sure predictions, unless they are used to calculate other coefficients of probabilities which would be small enough for application of *the single law of chance.*

For instance, we said that the probability that a chromosome S of A will be found in his brother B is 1/2, while it is only 1/8 that this chromosome of A will be found in his first cousin C. It may very well happen, however, that S is not found in B but is found in C, that is, there may be between the first cousins a resemblance or analogy which does not exist between the brothers.

But if we consider 100 pairs of brothers A_1, B_1; A_2, B_2; etc., and suppose that the hundred A's possess a given chromosome which determines in them a characteristic S, we may assert that this chromosome, and consequently this characteristic S, will be found an average of 50 times among the hundred B's. And, by virtue of the single law of chance, we conclude that it is impossible that S be found in all the hundred B's at the same time, or even in more than 95 of them, and equally impossible that S not be found in a single B, or even in only less than 5 of them. If instead of 100 pairs of brothers AB, we had considered 100 pairs of first cousins AC, the characteristic S would have been found on the average in 12.5 of them, and we could have asserted with certainty that it would not be found in more than 50 of them, while there would be an extremely small probability, although it would not be absolutely impossible, that it would not be found in any one of them.

If it is assumed a priori that we do not know whether the 100 pairs were formed of brothers or first cousins, but if we do know that the relation is the same for the 100 pairs, and the characteristic *S* is observed 60 times in both individuals, we can be certain they are brothers, while if *S* is found only 5 or 6 times, they are first cousins.

These examples suffice to show how the various results obtained in this chapter and the preceding can lead to sure predictions, when they are combined in such a way as to make applicable the single law of chance.

Note to Chapter Six: Since the original French edition of this work was published, it has been determined that there are 46 chromosomes (23 pairs of chromosomes) in the human species.

APPENDIX ONE

On Recurrences of Figures in Winning Lottery Numbers

1. Probabilities of the Various Types of Numbers. The problem of the probability of the recurrences of figures in winning lottery numbers, briefly mentioned in the first chapter, seems to deserve further development, for it is one of those problems which can contribute the most to the understanding of difficulties often encountered in numerical applications of the calculus of probabilities.

Let us consider all the six-figure numbers in the decimal system. They number one million, if we include numbers of less than six figures which may be completed at the left by zeros, and the number zero, which is written 000,000. These are the numbers which can be obtained by drawings out of six bowls arranged in a determined order, each containing the ten figures: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

We shall first calculate how many among these million six-figure numbers contain either 6 different figures or 5, 4, 3, 2, 1 different figures.

In a six-different-figure number, such as 324789 or 023586, the figure on the left may be any one of the ten figures, the following figure any one of the other nine figures, the third any one of the remaining eight, and so on up to the sixth figure, which is any one of the last five. There are then $10 \times 9 \times 8 \times 7 \times 6 \times 5 = 151{,}200$ numbers of 6 different figures.

Let us pass on to numbers of five different figures. One and only one figure is repeated once. That is what poker players call a pair. The repeated figure may be any one of the ten figures and may be put in any two of the six possible places, which gives fifteen possibilities[1] for each one of the figures, or in all, 150.

[1] The first figure may be put in any one of the six places and the second in any one of the other 5 places, which seems to give 30 possibilities. But if the second place is picked first and then the fourth, the distribution is the same as if the fourth were picked first and then the second. We must divide 30 by 2, which gives 15. This result may be verified by a direct listing of the 15 possible distributions.

When the repeated figure is put, for instance, in the second place and in the fifth, the number may be written

x3xx3x,

x designating the undetermined figures which must not be 3's.

The rightmost x may be replaced by any one of the 9 figures other than 3, the following x by any one of the 8 figures remaining, the other x's by any one of the 7 and any one of the 6 figures that remain. The number of possibilities is equal to

$$150 \times 9 \times 8 \times 7 \times 6 = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 3,$$

the triple of the number 151,200 of numbers formed with 6 different figures.

The number of possibilities with 5 different figures (and one pair) is therefore

$$151{,}200 \times 3 = 453{,}600.$$

Analogous reasonings determine the number of possibilities with 4 different figures. They may be divided into two categories, one made up of numbers with two pairs, such as 121472 or 003347, the other of numbers containing a figure repeated three times (a triplet), such as 303483. The first category (2 pairs) comprises 226,800 possibilities and the second (1 triplet) 100,800 possibilities, in all, 327,600 numbers with only 4 figures.

To obtain all the numbers containing 2 pairs, the two repeated figures must be picked first. The choice may be made in $\frac{10 \times 9}{2} = 45$ different ways. Any one of the ten figures may be picked, then any one of the remaining nine, in all, $10 \times 9 = 00$ choices. But each couple of two figures, such as 7 and 5, is obtained twice, since one may pick first 7 and then 5, or first 5 and then 7. The number of couples of two figures is therefore half of 90, or 45. Let 7 and 5 be the chosen couple: 7 may be put in any one of the six places, then in any one of the remaining 5 places. The total number of choices is 6×5, but this number must be divided by 2 for the reason just given. There are 15 ways to choose the places of the two 7's. When the two 7's are placed, there remain four vacant places, and there are six ways to place the two 5's. When the two 7's and two 5's are placed, the distribution may be

x577x5,

where the first x may be replaced by any one of the 8 other figures and the second x by any one of the 7 remaining figures. We obtain finally a number of possibilities equal to

$$45 \times 15 \times 6 \times 8 \times 7 = 5 \times 9 \times 8 \times 7 \times 6 \times 5 \times 3.$$

This number is therefore half of the already calculated number

$$10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 3 = 453{,}600.$$

It is equal to 226,800.

It is a rather remarkable fact that there is exactly the same total number of pairs in the one-pair and the two-pair numbers. This fact does not occur for all the values of the total number of figures involved (equal to 10 here, since the decimal system is used) and of the number of figures forming the considered possibilities.[2]

To obtain all the numbers containing a triplet, the triplet must first be picked out, which may be done in ten different ways. The three places which it occupies may then be chosen in $\frac{6 \times 5 \times 4}{1 \times 2 \times 3} = 20$ different ways. There are 200 distributions like the following

x88xx8,

each of which may be completed in $9 \times 8 \times 7$ different ways by three figures different from 8 and differing from each other. There are in all $200 \times 9 \times 8 \times 7 = 100{,}800$ numbers containing a triplet.

Let us pass on to numbers containing only 3 different figures. They may contain 3 pairs, such as 422477 (10,800 of these) or a pair and a triplet, such as 422274 (43,200 of these), or finally a quadruplet, such as 447484 (10,800 of these): in all, 64,800 numbers formed by 3 figures.

[2] It can easily be shown that if the total number of figures involved is $n = 2K^2 \pm K$, K being any integral number, this property subsists if the number of figures appearing in a numter, is $p = 2K + 2$ if $n = 2K^2 + K$ and $p = 2K + 1$ if $n = 2K^2 - K$. The above result is for $K = 2$, $n = 2K^2 + K = 10$, $p = 2K + 2 = 6$.

Without going into the details of reasonings always based on the same principles, let us indicate how the preceding numbers are obtained.

Numbers containing three pairs:

$$\frac{10 \times 9 \times 8}{1 \times 2 \times 3} \times \frac{6 \times 5}{1 \times 2} \times \frac{4 \times 3}{1 \times 2} = 10{,}800.$$

Numbers containing a pair, a triplet and another figure:

$$\frac{10 \times 9 \times 8 \times 6 \times 5 \times 4}{2} = 43{,}200.$$

Numbers containing a quadruplet (a figure occurring 4 times) and two other different figures:

$$10 \times \frac{6 \times 5 \times 4 \times 3}{1 \times 2 \times 3 \times 4} \times 9 \times 8 = 10{,}800.$$

We have calculated (p. 11) the number of combinations containing only two different figures. They fall into three categories.

Numbers containing a quintuplet (a figure occurring 5 times) and another figure:

$$10 \times 9 \times 6 = 540.$$

Numbers containing a quadruplet and a pair:

$$10 \times 9 \times \frac{6 \times 5}{1 \times 2} = 1{,}350.$$

Numbers containing two triplets:

$$\frac{10 \times 9}{1 \times 2} \times \frac{6 \times 5 \times 4}{1 \times 2 \times 3} = 900.$$

Finally, numbers formed by means of a single figure (including 000000, but leaving aside 333, for instance, which must be written 000333): 10.

These results may be gathered as in the following table.

TABLE I

Number of different figures	*Example*	*Number corresponding to each example*	*Total number for each number of different figures*
6	327689	151,200	151,200
5	327683	453,600	453,600
4	327376 327336	226,800 100,800	327,600
3	071701 007017 723777	10,800 43,200 10,800	64,800
2	556555 556565 556566	540 1,350 900	2,790
1	333333 or 000000	10	10
TOTAL		1,000,000	1,000,000

Many readers will certainly be surprised at these results, which are, however, incontestable. Since there are ten figures and only six bowls, one could have expected the most frequent case to be that in which every bowl would give a different figure. This happens only about 15 times in 100, whereas more than 45 times in 100 the same figure is drawn twice; and nearly 33 times in 100 only 4 different figures are drawn, two of them being drawn twice each (nearly 23 times in 100) or one figure being drawn 3 times (about 10 times in 100).

If attention is fixed on a single drawing of the lottery, it will often happen that the proportions of winning tickets with, respectively, 6, 5, 4, 3 different figures, will be quite different from those which have just been calculated. But if we consider a number of drawings large enough to involve at least a hundred, or preferably, several hundred important prizes, the proportions will be seen to approach closely those given by our table. It will be seen in particular that by far the most important case, providing nearly half of the winning tickets, is that of numbers in which one, and only one, figure occurs twice. Of course, these breakdowns must not overlook the zeros which must be written at the left, so that all the numbers have exactly six figures.

2. Results Relative to the Repetitions of a Particular Figure. It is interesting to compare the results which we have just obtained with those obtained by fixing attention on a particular figure, 7, for instance, and classifying the numbers according to the recurrence of 7.

Numbers not containing 7.—Each one of the six figures of these numbers may be chosen arbitrarily among the other nine figures. The number of combinations is:

$$9 \times 9 \times 9 \times 9 \times 9 \times 9 = 9^6 = 531{,}441.$$

Numbers containing one, and only one, 7.—The figure 7 may be put in any one of the six places, and then in any one of the 5 remaining places any one of the other nine figures. The number of combinations is:

$$6 \times 9 \times 9 \times 9 \times 9 \times 9 = 6 \times 9^5 = 354{,}294.$$

Numbers containing two, and only two, 7's. The two 7's may be placed in $\frac{6 \times 5}{1 \times 2} = 15$ different ways, and one of the other 9 figures may be written in each one of the other 4 places. The number of combinations is

$$15 \times 9 \times 9 \times 9 \times 9 = 15 \times 9^4 = 98{,}415.$$

Numbers containing three, and only three, 7's. The three 7's may be placed in $\frac{6 \times 5 \times 4}{1 \times 2 \times 3} = 20$ different ways, and we have in all:

$$20 \times 9 \times 9 \times 9 = 20 \times 9^3 = 14{,}580 \text{ combinations.}$$

Numbers containing four, and only four, 7's.—There are fifteen possible places for the four 7's, and

$$15 \times 9 \times 9 = 1{,}215 \text{ combinations.}$$

Numbers containing five, and only five, 7's.—There are

$$6 \times 9 = 54 \text{ combinations.}$$

Finally, there is only one number 777777 of six 7's.

Let us summarize in a table the results obtained.

TABLE II

Number of 7's	*Number of combinations*
0	531,441
1	354,294
2	98,415
3	14,580
4	1,215
5	54
6	1
TOTAL	1,000,000

It may be observed that the numbers obtained are the terms of the development of the sixth power of the binomial 9 + 1:

$$(9 + 1)^6 = 9^6 + 6 \times 9^5 + 15 \times 9^4 + 20 \times 9^3 \ldots$$

Table II suggests several interesting remarks.

It may first be noted that more than half of the combinations (531,441 in a million) do not contain 7. Suppose six drawings are made in each one of which the probability of drawing 7 is one tenth. The sum of these probabilities is six tenths, that is, more than one half. This shows that the probabilities must not be added indiscriminately. What may be added are the mathematical expectations, that is, the probabilities of gain of a gambler who bets on the 7. If this gambler puts up one franc, he should, in fairness, be given 10 francs when the 7 comes out. If six drawings are made at once, the gambler must put up six francs and will receive as many times ten francs as the number 7 comes out.

The table shows that the gambler has about 53 chances in 100 to lose his 6 francs, a little more than 35 chances in 100 to win 10 francs, nearly 10 chances in 100 to win 20 francs, 14 chances in 1,000 to win 30 francs, about 1 chance in 1,000 to win 40 francs. These possibilities of relatively high winnings compensate for the fact that he loses his wager of 6 francs more than one time in two.

Let us now consider the case where 7 appears more than once. In a million trials, there are 98,415 pairs of 7, 14,580 triplets of 7, 1,215 quadruplets of 7 and 54 quintuplets of 7.

Since one may reason regarding each one of the ten figures exactly as regarding 7, there are among the million combinations 984,150 pairs, nearly a million. It would be wrong to conclude that almost all the numbers contain a pair. Table I tells us that only 453,600 numbers contain one, and only one, pair. To obtain the total number of pairs, account must be taken of the combinations with 2 or 3 pairs and of those where the pair is accompanied by a triplet or a quadruplet. Table I gives us:

		Pairs
453,600	numbers with one isolated pair, or	453,600
226,800	,, ,, two pairs, or	453,600
10,800	,, ,, three pairs, or	32,400
43,200	,, ,, a pair and a triplet, or	43,200
1,350	,, ,, a pair and a quadruplet, or	1,350
735,750	numbers containing in all	984,150

The result agrees with what was deduced from Table II, confirming the exactness of our calculations.

Table II also shows that there are in all 145,800 triplets, which, according to Table I, are thus distributed:

100,800	numbers with one isolated triplet	100,800
43,200	,, where the triplet is accompanied by a pair	43,200
900	,, with two triplets	1,800
	TOTAL	145,800

Finally, there are, according to Table II, 12,150 quadruplets in all, of which 10,800 are isolated and 1,350 are accompanied by a pair, according to Table I.

Let us indicate, in conclusion, a curious consequence of the figures of Table II.

Let us suppose that a player bets on the drawing of pairs and is promised as many times 10 francs as the number of pairs in the drawn combination. If he plays a million times and all the tickets are drawn, he will see 984,150 pairs in all, or nearly one million. If his wager is 10 francs, the game is about equitable; it only holds back a profit of 15 to 16 per 1,000 or about $1\frac{1}{2}$ per cent for the entrepreneur of the lottery, who has contracted to pay 10 francs for every pair drawn. But it is worthy of notice that the game can be made entirely equitable by agreeing that every triplet, quadruplet, quintuplet or sextuplet will pay, not 10 francs as does the pair, but only one franc. According to Table I, the total number of triplets, quadruplets, quintuplets and sextuplets for the figure 7 is:

$$14,580 + 1,215 + 54 + 1 = 15,850.$$

The total number would be 10 times as great for all the figures together, but if only one tenth of the wager (1 franc instead of 10) is put up, we must simply add 15,850 to 984,150, which makes exactly one million.

This rather remarkable result is a consequence of the following relation, which our readers will easily verify:

$$150 \times 9^4 + (10^6 - 9^6 - 6 \times 9^5 - 15 \times 9^4) = 10^6.$$

Thus, the following game is perfectly equitable: Peter gives Paul 10 francs before the drawing of the lottery, and if the winning number contains pairs, Paul pays to Peter as many times 10 francs as there are pairs. If instead of pairs, or in addition to them a figure occurs more than twice,[3] Paul pays one franc to Peter for each one of these groups of more than two identical figures (triplets, quadruplets, quintuplets, or sextuplets). Thus Peter's possible gains are the following (from which should be deducted his 10 franc wager):

One pair ..	.. 10 francs	Triplet ..	.. 1 franc
Two pairs ..	.. 20 ,,	Two triplets ..	.. 2 francs
Three pairs ..	.. 30 ,,	Quadruplet ..	.. 1 franc
Pair and triplet	.. 11 ,,	Quintuplet ..	.. 1 ,,
Pair and quadruplet	.. 11 ,,	Sextuplet ..	.. 1 ,,

[3] If a figure occurs exactly three times it must be regarded as a triplet; if exactly four times, then a quadruplet and not two pairs; if exactly five times, then a quintuplet and not a pair and triplet; etc.—PUBLISHER'S NOTE.

APPENDIX TWO

On Poisson's Formula

1. Poisson's Formula. Poisson's formula reveals the probabilities relative to the intervals of time which separate fortuitous events following each other without any other law than the recognized existence of a certain *average frequency*. If, for instance, a roulette wheel functions at the rate of one play per minute, each number, say, 17, will come up, on the average, once every 37 minutes: this is the average frequency.[1] There are some very important phenomena which fall under this definition. Such is the case of emissions of particles which correspond to the disintegration of certain radioactive molecules. For a given mass of radium, the average number of disintegrated atoms in a given interval of time is a well-determined constant.

The time intervals may be represented by proportional segments of a straight line. Instead of the distribution of the events in time, one may speak of the distribution of points on the line. These points may be considered as distributed at random, under the sole condition that their *mean density* is constant, the density being the number of points per unit of length; if it is designated by d, the number of points situated on a segment a will be, *on the average*, ad.

Let us then consider either an interval of time or a fixed segment of the straight line, and let us designate by $b = ad$ the average number of events or points which may be expected to be observed in the given interval of time or on the given segment of the line. This number b is generally not an integer. Even when b is an integer, one may not always observe the precise number b. Poisson's law shows the probability that n events (or n points) will be precisely observed, instead of the average

[1] Actually, Poisson's law is a limit law which would apply rigorously only if a roulette wheel could be imagined whose cadence was more and more rapid, while the possible numbers increased in quantity. For a wheel of 600 numbers, say, playing once per second, every number would come up *on the average* every ten minutes.

number b. This probability P is:

$$P = e^{-b}\frac{b^n}{n!} \qquad (1)$$

Such is Poisson's formula, in which e designates, as usual, the base of Napier's logarithms ($e = 2.718281828\ldots$).

If $b = 1$, formula (1) becomes

$$P = \frac{1}{e}\frac{1}{n!}. \qquad (2)$$

It is by formula (2) that the results given in Chapter Four were calculated.

We have

$$\frac{1}{e} = 0.36788\ldots.$$

and formulas (1) and (2) apply also when $n = 0$, if $n!$ is replaced by 1.

($n!$ designates the product of the first n integers. $(n + 1)! = (n + 1)n!$, and if $n = 0$, then $0! = 1$.)

2. The Problem of Waiting at the Window. It is thanks to Poisson's formula (and to other calculations) that we were able to obtain the results indicated in Chapter Four relative to the problem of the waiting at a window. Let us say that the number of clients at the window is N per day, and that each client remains a minutes. We suppose that the product Na is less than the total time the office is open, or more precisely, that $Na = D\sigma$, σ being a number less than 1.

As we have remarked, all clients together may be divided into series, each series being composed of clients succeeding one another without interruption, while in the interval between two series the window remains free. The probability that a series be composed of n clients is given by the formula.[2]

$$P_n = e^{-n\sigma}\sigma^{n-1}\frac{n^{n-2}}{(n-1)!}.$$

This formula gave the numerical results cited in Chapter Four.

[2] For a proof, see Émile Borel, *Sur l'emploi du théorème de Bernoulli, pour le calcul d'une infinité de coefficients. Application au problème d'attente à un guichet* (Comptes rendus de l'Académie des Sciences, March 1942).

APPENDIX THREE

On Mortality Tables and the Statistics Relative to Causes of Death

1. Mortality Tables of French Companies. It was in the middle of the eighteenth century that Deparcieux established the first Mortality Tables and calculated the probabilities of death which derive from them. Since then statistical researches on mortality have been very numerous and highly perfected, particularly in certain countries and large cities. The statistics gathered by insurance companies on their own clientele are more restricted, but, from a certain point of view, they are at least as interesting as the more general statistics.

As a sample of the tables so established for insurance companies, we reproduce, from the *Annuaire du Bureau des Longitudes* of 1937, Table F.A. (French annuitants) and Table F.I. (French insurants) of the French companies (1895).

TABLE F.A. (FRENCH ANNUITANTS) OF FRENCH COMPANIES (1895)

Ages	*Living*	*Ages*	*Living*	*Ages*	*Living*	*Ages*	*Living*
0	1,000,000	18	835,173	36	740,070	54	613,494
1	963,985	19	829,762	37	734,545	55	603,634
2	937,488	20	824,159	38	728,922	56	593,302
3	917,939	21	818,471	39	723,190	57	582,465
4	903,486	22	812,809	40	717,338	58	571,092
5	892,765	23	807,271	41	711,352	59	559,149
6	884,754	24	801,926	42	705,219	60	546,604
7	878,676	25	796,786	43	698,925	61	533,427
8	873,932	26	791,817	44	692,452	62	519,588
9	870,056	27	786,827	45	685,784	63	505,060
10	866,684	28	781,811	46	678,902	64	489,820
11	863,529	29	776,764	47	671,787	65	473,851
12	860,371	30	771,681	48	664,417	66	457,139
13	857,043	31	766,556	49	656,770	67	439,680
14	853,426	32	761,383	50	648,823	68	421,478
15	849,446	33	756,156	51	640,548	69	402,549
16	845,069	34	750,866	52	631,921	70	382,919
17	840,298	35	745,508	53	622,913	71	362,630

Ages	*Living*	*Ages*	*Living*	*Ages*	*Living*	*Ages*	*Living*
72	341,741	81	145,553	90	20,791	99	225
73	320,328	82	125,891	91	14,874	100	103
74	298,484	83	107,374	92	10,296	101	44
75	276,325	84	90,185	93	6,873	102	17
76	253,984	85	74,477	94	4,408	103	6
77	231,618	86	60,372	95	2,706	104	2
78	209,398	87	47,947	96	1,583	105	1
79	187,512	88	37,232	97	878		
80	166,162	89	28,204	98	459		

TABLE F.I. (FRENCH INSURANTS) OF FRENCH COMPANIES (1895)

Ages	*Living*	*Ages*	*Living*	*Ages*	*Living*	*Ages*	*Living*
0	1,000,000	27	786,713	54	584,594	81	91,047
1	963,985	28	781,578	55	572,246	82	76,094
2	937,488	29	776,368	56	559,322	83	62,588
3	917,939	30	771,075	57	545,797	84	50,588
4	903,486	31	765,690	58	531,649	85	40,118
5	892,765	32	760,203	59	516,861	86	31,159
6	884,754	33	754,606	60	501,417	87	23,658
7	878,676	34	748,887	61	485,307	88	17,523
8	873,932	35	743,036	62	468,525	89	12,632
9	870,056	36	737,039	63	451,075	90	8,841
10	866,684	37	730,884	64	432,964	91	5,992
11	863,529	38	724,556	65	414,214	92	3,920
12	860,371	39	718,042	66	394,851	93	2,468
13	857,043	40	711,324	67	374,918	94	1,490
14	853,426	41	704,386	68	354,468	95	859
15	849,446	42	697,210	69	333,567	96	471
16	845,069	43	689,777	70	312,299	97	245
17	840,298	44	682,067	71	290,759	98	120
18	835,173	45	674,058	72	269,062	99	55
19	829,762	46	665,729	73	247,333	100	23
20	824,159	47	657,056	74	225,714	101	9
21	818,471	48	648,015	75	204,359	102	3
22	812,809	49	638,581	76	183,430	103	1
23	807,271	50	628,727	77	163,096		
24	801,926	51	618,429	78	143,530		
25	796,786	52	607,659	79	124,896		
26	791,780	53	596,389	80	107,354		

2. Tables of General Statistics of France. Other tables are also to be found in the *Annuaire du Bureau des Longitudes*, especially those established every ten years in the general statistics of France, by combining the mortality according to official records over a six-year period with a census taken about the middle of this period, a census showing the distribution of the population according to ages.

To save space, these tables have been simplified by omitting superfluous decimals and restricting the figures, after the first five years, to multiples of five years of age.

The table shows the number of survivors in 1,000 births, the yearly quotient of mortality at every age (that is, the average number of deaths in the year for 1,000 persons of that age), and finally the life expectancy, that is, the mathematical expectancy of a gambler who would collect 1 franc per year of life of a given person (leaving aside, of course, the interest on the money). The tables are computed separately for males and females. The quotients of mortality, for almost all the ages, are markedly higher for men than for women. These mortality quotients refer to 1,000 inhabitants.

We deduce from these tables that for men of 60, for instance, the number of survivors in 1,000 births is 544, the mortality quotient is 29 over 1,000 and the life expectancy 13.8 years, or a little more than 13 years 9 months and a half. The average age to be reached by a rather large number of men of 60 is therefore 73 years and 9 months.

MORTALITY TABLES
OF THE FRENCH POPULATION (1928–1933)

A = ages.
SM = survivors (male)
SF = survivors (female)
MQ = mortality quotient for 1,000 (males) from the age A to $A + 1$.
FQ = mortality quotient for 1,000 (females) from the age A to $A + 1$.
EM = life expectancy (male).
EF = life expectancy (female).

A	SM	MQ	EM	SF	FQ	EF	A
0	1,000	90	54.3	1,000	72	59.0	0
1	910	17	58.6	928	15	62.5	1
2	894	6.7	58.6	914	6.3	62.5	2

A	SM	MQ	EM	SF	FQ	EF	A
3	888	4.3	58.0	909	4.0	61.9	3
4	885	3.4	57.3	905	3.2	61.1	4
5	882	2.8	56.5	902	2.8	60.3	5
10	872	1.6	52.1	892	1.6	55.9	10
15	864	2.5	47.5	884	3.0	51.4	15
20	849	5.2	43.3	867	4.8	47.4	20
25	827	5.2	39.4	846	5.0	43.5	25
30	805	5.9	35.4	825	4.8	39.5	30
35	780	7.1	31.5	806	5.1	35.4	35
40	750	8.9	27.6	784	6.1	31.4	40
45	713	12	23.9	759	7.5	27.3	45
50	669	15	20.3	727	9.8	23.4	50
55	613	21	16.9	688	13	19.6	55
60	544	29	13.8	637	19	15.9	60
65	458	42	10.9	567	30	12.6	65
70	354	64	8.3	472	48	9.6	70
75	261	92	6.5	375	72	7.5	75
80	125	153	4.4	120	128	5.1	80
85	45	234	3.2	91	200	3.6	85
90	9.6	303	2.6	24	285	2.7	90
95	1.4	334	2.3	3.8	336	2.4	95
100	0.2	348	1.5	0.5	346	2.1	100

To help the reader realize the progressive decrease in mortality, we give, for every ten years, the number of survivors for 1,000 births, according to the General Statistics of France for the six-year periods ending in 1903 and 1913, the four-year period ending in 1923, and the five-year period ending in 1938.

A considerable constant and regular amelioration is to be observed in the number of survivors at all ages.

SURVIVORS IN 1,000 BIRTHS

FOR THE PERIODS ENDING

IN 1903 (1898–1903), IN 1913 (1908–1913), 1923 (1920–1923) and 1938 (1934–1938) (*General Statistics of France*).

	Males				*Females*			
Ages	1903	1913	1923	1938	1903	1913	1923	1938
1	837	866	892	924	864	888	912	940
10	759	806	845	891	786	827	866	911
20	729	779	819	872	752	797	837	891
30	677	727	767	831	701	750	790	843
40	616	666	712	775	646	698	742	819
50	538	583	638	691	584	636	684	763

	Males				*Females*			
Ages	1903	1913	1923	1938	1903	1913	1923	1938
60	432	465	521	561	494	545	595	672
70	275	295	344	366	341	383	436	508
80	88	97	119	135	128	150	182	241
90	7.3	7.6	8.8	12	15	17	20	35

After 1938 the number of deaths increased considerably during the years of war and occupation, and decreased greatly in the years 1946, 1947, 1948. On the other hand, in the first quarter of 1949, a severe epidemic of grippe caused 198,000 deaths, 57,000 more than in 1948. It will not be possible for several years to know whether the decrease in mortality is an important and lasting phenomenon, resulting from a proven amelioration due to the progress of medicine and hygiene.

The following table gives the numbers of deaths from 1938 to 1948, according as the deaths from war casualties are included or not.

A and B: number of deaths, by thousands of inhabitants, from 1938 to 1948.
A (including the officially recorded deaths from war casualties).
B (not including deaths from war casualties).
A′ and B′ (respective percentages per 10,000 inhabitants).

	'38	'39	'40	'41	'42	'43	'44	'45	'46	'47	'48
A	647	642	760	674	657	631	744	658	542	533	506
A′	154	153	185	170	167	161	191	166	134	130	122
B	647	632	738	673	654	624	664	656	542	533	506
B′	154	151	180	170	166	160	170	165	134	130	122

3. Tables of Survival by Generations. In the preceding tables the percentage of mortality for any one age is computed for a given year by dividing the number of officially recorded deaths for that age by the total number of persons of that age, as may be determined from the census. These percentages are naturally adjusted so that the curve representing them seems to be continuous, giving the number of survivors at every age.

In the Tables of Survival by generations given below, following Pierre Delaporte (publications of the General Statistics), we proceed differently. For the generation born in 1820, for instance, the mortality percentage at the age of 55 is computed from the statistics of the year 1875. That is why these tables, computed at a time when the known statistics did not go beyond the year 1935, stop at age 75 for the generation born in 1860 and at age 35 for the generation born in 1900.

Here too some adjustments are, of course, necessary to bring out the true general behavior of the phenomena, a behavior which would be perturbed by accidental causes occurring in such or such a year.

MORTALITY RATE PER GENERATION

(DEATHS FROM AGE n TO $n + 1$ PER 100,000 OF AGE n)

Ages	*Born in* 1820		*Born in* 1860		*Born in* 1900	
n	*Males*	*Females*	*Males*	*Females*	*Males*	*Females*
0	17,600	15,270	16,500	16,500	16,500	13,600
1	6,400	6,200	7,470	6,780	3,000	3,170
5	1,520	1,500	1,350	1,330	530	570
10	635	735	500	620	240	260
15	575	716	450	570	280	370
20	872	878	820	720	620	520
25	1,020	926	960	790	590	550
30	825	964	830	840	600	480
35	930	1,004	960	870	680	460
40	1,074	1,068	1,110	890		
45	1,310	1,177	1,350	970		
50	1,545	1,382	1,610	1,150		
55	2,090	1,750	2,090	1,460		
60	3,000	2,560	3,000	2,080		
65	4,460	3,860	4,460	3,160		
70	6,720	5,900	6,720	4,940		
75	10,620	9,200	10,620	7,830		
80	16,050	14,000				
85	22,800	20,500				
90	29,300	26,600				
95	35,500	33,600				
99	42,000	39,900				

TABLES OF SURVIVAL OF GENERATIONS FOR 100,000 BIRTHS

	Born in 1820		*Born in* 1860		*Born in* 1900	
Ages	*Males*	*Females*	*Males*	*Females*	*Males*	*Females*
1	82,400	84,730	83,500	83,500	83,500	86,400
5	70,948	73,167	71,192	71,879	77,993	80,675
10	67,133	69,252	68,168	68,584	76,561	79,066

Ages	Born in 1820		Born in 1860		Born in 1900	
	Males	*Females*	*Males*	*Females*	*Males*	*Females*
15	65,268	66,834	66,694	66,692	75,677	78,012
20	63,089	64,230	64,828	64,572	74,086	76,311
25	59,853	61,382	61,794	62,211	71,810	74,280
30	57,206	58,552	59,140	59,732	69,731	72,340
35	54,771	55,738	56,601	57,230	67,575	70,670
40	52,142	52,926	53,778	54,773		
45	49,164	50,074	50,628	52,315		
50	45,829	47,016	47,038	49,670		
55	41,955	43,574	43,028	46,619		
60	37,157	39,327	38,106	42,858		
65	31,003	33,708	31,795	37,851		
70	23,640	26,693	24,244	31,261		
75	15,464	18,478	15,859	22,983		
80	7,838	10,336				
85	2,785	4,200				
90	636	1,150				
95	94	200				
100	8	20				

Only ten or fifteen years hence, when these Tables of Survival by generations have been extended, will it be possible to determine whether the net amelioration which appears at the beginning of the tables holds up till the end. In other words, it is not yet possible to know for certain whether the decrease in mortality during childhood, adolescence and even maturity continues in old age, causing a real prolongation of life, or whether this mortality, smaller before the age of 50 or 60, is not in some way compensated by a greater mortality after this age.

4. Statistics of Deaths According to their Causes. As we have said, statistical science has made great progress in the last ten years. We give below the new official classification of the causes, restricting ourselves to the main categories (each one of which is broken down in a more or less extended list of more precise designations).

NOMENCLATURE OF THE CAUSES OF DEATH

I. Infections and parasitic diseases.
II. Cancers and other tumors.
III. General diseases and avitaminosis.

IV. Diseases of the blood and hematopoietic organs.
V. Chronic poisonings and intoxications.
VI. Diseases of the nervous system and sensory organs.
VII. Diseases of the circulatory system.
VIII. Diseases of the respiratory system (except tuberculosis, included in I).
IX. Diseases of the digestive system.
X. Diseases of the urinary system and the genital system.
XI. Diseases of pregnancy, delivery and puerperal state.
XII. Diseases of the skin and cellular tissue.
XIII. Diseases of the bones and organs of motion.
XIV. Congenital defects of conformation (stillborn children not included).
XV. Diseases peculiar to the first year of life (stillborn children not included).
XVI. Senility, old age.
XVII. Violent or accidental deaths.
XVIII. Undetermined causes.
XIX. Grand total.

The causes III, IV, V, XI, XII, XIII, XIV, which produce a relatively small number of deaths are grouped below under the letter G.

DEATHS ACCORDING TO SEX IN 1948

Causes	*Males*	*Females*	*Total*
I	24,499	16,564	41,063
II	34,352	36,512	70,864
VI	31,812	36,195	68,007
VII	49,609	50,810	100,429
VIII	24,747	23,266	48,013
IX	14,561	11,605	26,166
X	14,367	9,342	23,709
XV	8,132	5,892	14,024
XVI	14,788	23,714	38,502
XVII	18,872	7,536	26,408
G	6,709	7,081	13,790
XVIII	18,539	16,773	35,312
XIX	260,987	245,290	506,287

DEATHS ACCORDING TO AGE IN 1948
FROM THE CHIEF CAUSES (MALES)

Ages	*I*	*II*	*VI*	*VII*	*VIII*	*IX*	*XVII*
0–1	1,813	26	2,686	202	4,555	3,735	313
1–4	872	61	522	51	728	445	581
5–9	288	57	141	78	67	70	315
10–14	205	51	117	90	46	84	309
15–19	563	73	153	139	69	107	811
20–24	1,301	117	167	191	132	147	1,241
25-29	1,596	147	178	231	141	176	1,171
30–34	1,075	127	162	225	122	164	926
35-39	1,818	407	325	537	256	329	1,434
40–44	2,246	1,033	503	996	488	573	1,763
45-49	2,672	2,146	855	1,724	851	894	1,768
50-54	2,143	2,831	1,203	2,339	1,009	1,024	1,481
55-59	1,965	3,525	1,695	3,270	1,190	1,029	1,230
60-64	1,840	4,745	2,692	5,265	1,749	1,308	1,293
65-69	1,625	5,600	4,027	6,962	2,279	1,352	1,246
70-79	2,000	10,579	11,098	17,977	6,719	2,269	2,131
80–99	477	2,827	5,288	9,331	4,346	855	859

DEATHS ACCORDING TO AGE IN 1948
FROM THE CHIEF CAUSES (FEMALES)

Ages	*I*	*II*	*VI*	*VII*	*VIII*	*IX*	*XVII*
0–1	1,456	24	1,790	151	3,400	2,628	205
1–4	813	53	408	55	663	325	433
5–9	299	49	107	57	67	50	152
10–14	250	34	64	68	52	64	73
15-19	723	60	126	113	93	91	200
20–24	1,358	93	128	189	107	123	256
25–29	1,483	118	139	283	151	135	236
30–34	859	205	132	264	123	127	229
35–39	1,093	641	260	474	191	244	323
40–44	1,063	1,127	392	646	228	418	365
45–49	955	1,955	745	1,075	354	570	386
50–54	838	2,728	1,259	1,513	466	670	468
55–59	914	3,513	1,904	2,250	672	720	451
60–64	992	4,501	2,948	3,794	1,116	898	510
65–69	1,015	5,403	4,442	6,083	1,757	1,065	563
70–79	1,702	11,261	13,021	18,963	6,775	2,194	1,353
80–99	751	4,747	8,380	14,832	7,051	1,283	1,333

We do not reproduce statistics by provinces or past statistics. The readers who are interested in them will find them in the *Annuaire de la Statistique générale de la France* published by the Service de la Statistique générale de la France, 11 Boulevard Haussmann, Paris (9e), and printed by the Imprimerie Nationale.

INDEX

(continued from front flap)

Science, Theory and Man, Erwin C. Schrödinger. $1.35

Elementary Metallurgy and Metallography, Arthur M. Shrager. $2.25

Great Ideas and Theories of Modern Cosmology, Jagjit Singh. $2.00

Great Ideas of Information Theory, Language and Cybernetics, Jagjit Singh. $2.00

Great Ideas of Modern Mathematics: Their Nature and Use, Jagjit Singh. $2.00

The Architecture of the Body, C. U. M. Smith. Clothbound $6.00

Bridges and Their Builders, David B. Steinman and Sara R. Watson. $2.00

Mathematics in Action, O. G. Sutton. Clothbound $3.50

Physics, The Pioneer Science, Lloyd W. Taylor. Two-volume set $4.50

An Introduction to Patents for Inventors and Engineers, C. D. Tuska. $1.50

Basic Theory and Application of Transistors, U. S. Department of the Army. $1.25

PROBABILITIES AND LIFE

Émile Borel

What is the probability if you are now thirty-five years old that you will live to be 100? How likely is it that next Tuesday will be a clear day? What are your chances of catching a bus within one minute after reaching the bus stop, if the buses arrive alternately at 10 and 30 minute intervals? How likely is it that two first cousins (who have the same maternal grandmother) will have, among their chromosomes, one chromosome in common?

You may have encountered similar problems at one time or another, but perhaps you weren't aware that they can all be answered by means of the mathematics of probability, which computes the probabilities of complex matters through the known probabilities of simpler ones. This book — by Émile Borel, one of the leading French mathematicians of the past hundred years — makes use of certain results of the mathematics of probabilities to solve a number of problems that directly concern every man — problems that, for the most part, are related either to everyday living or to illness and death: computation of life expectancy tables, chances of recovery from various diseases, probabilities of job accidents, weather predictions, games of chance, and other such matters. The emphasis throughout is on the results, rather than the actual processes, of the mathematics of probability, though some indication of the mathematical proof is given in order to show what one can accomplish in this field.

Beginning with a discussion of the connection between mathematical probability and the psychology of gamblers, Borel takes up the probabilities of life and death, and the difficulties we encounter in trying to think rationally about them. The next two chapters are concerned with negligible probabilities on various scales (human, terrestrial, cosmic, supercosmic), and with others that are very small but not entirely negligible. The last two chapters investigate the probabilities of illnesses and accidents, and some curious applications of mathematical probability to heredity in the human race.

This book differs from others of its kind in that it does not attempt to cover the entire field of mathematical probability — it omits probability in scientific research, for example — and concentrates instead on probability in everyday situations. Simple in style and free of technical terminology, the book is entirely comprehensible to the layman, and provides fine reading material for those entering on the study of probability.

New English translation. Index, 3 Appendixes of statistical studies and tables. vi + 87pp. 5⅜ x 8½.
T121 Paperbound

A DOVER EDITION DESIGNED FOR YEARS OF

We have made every effort to make this the be[...] with minimal show-through; it will not discolor [...] sewn in signatures, in the method traditionally [...] drop out, as often happens with paperbacks he[ld together with glue. Pages open flat] for easy reference. The binding will not crack [or split. This is a permanent book.]

DOVER