THE ANNALS OF THE HITOTSUBASHI ACADEMY

Vol. IX No. 2 April 1959

THE LAW OF GREAT NUMBERS AND THE PRINCIPLES OF PROBABILITY*

By KOTARO FUJIMOTO

Professor Emeritus, Hitotsubashi University

I

Statistics is a science making mass phenomena as its object. In Germany, until recent times, it was asserted that mass observation was its major object of inquiring into the natures of different phenomena in human society, e. g., in population, economy or culture.

Accordingly, it has been said that in the case of the natural sciences, e. g., meteorology, statistics merely fulfils the rôle of a purely supplementary measure, however, in the sphere of economics or demography, as the various facts and their relations could be made explicit mainly through statistics, it has had to be taken as a science.

The reason why Prof. G. v. Mayr gave the title to his book "Statistics and Sociology,"[1] is because he thought this point should be emphasized.

But during the past thirty years or so after his death (died 1925), statistics has made a remarkable progress in its analytical methods, especially in the last

* I wrote this essay with the intention of making the students, attending my lecture, understand the applications and the limits of the law of great numbers, and I had no intention of helping scholars specializing in statistics. The theory of probability is naturally a principle in mathematics, closely connected with the law of great numbers, and statistics of today rests on the theory of probability.

I intended, therefore, to make the students understand the meaning of the theory of probability, which is an indispensable instrument for the students of social sciences, for example, economics, sociology and the like, but I did not intend to clarify how to deal mathematically with the theory of probability. Therefore, probability proper was intentionally shunned here, and left to the appropriate reference books in mathematics.

[1] G. v. Mayr, *Statistik und Gesellschaftslehre*, Tübingen 1914.

decade, and it can not be denied that the statistical methods have made great strides both in England and America.

The fact that the year the book written by R. A. Fisher,[2] happened to be published in the same year as that in which Prof. Mayr died, seems to suggest something significant.

Fisher's "Statistical Methods" was a product arising from actual necessities in his studies at the Rothamsted Biology Research Instutite, hence, at first, his statistical methods might not have been intended to have social science as its direct object, but soon after, these methods began to be used in inquiring into social phenomena; it was seen that research by using his methods could get much better results compared with those used in the past. By, such as the theory of the sampling method, significance test as well as degree of freedom, furthermore, correlation, discovered by Galton and Pearson, regression coefficient, etc., statistics seems to have been entirely renewed. Briefly, it may not be too much to say that they are the causes of the revolutional changes brought about in statistics.

Soon after the World War II, in Japan, Mr Deming, an American, made several visits to our country and introduced these theories and their applications in the field of statistics. Since then, the number of scholars intending to study statistics has remarkably increased in a short period, so that it is needless to say that such a popularity of the study of statistics, due to its progress in Japan, is unprecedented.

Accordingly, it is natural that there is a tendency in studying statistics, to consider it as a method or a methodology which has been in prevalence in England and America. Statistical methods, as shown above, have enabled us to analyse social phenomena more explicitly than before.

Recently, by such as quality control, market research, public opinion and different kinds of index numbers, statistics is no longer allowed to remain behind the cloistered world of abstract science, but in being used in industrial management, etc.; statistics may be looked upon as an indispensable instrument in our practical life. Formerly, it was thought that business risk was considerably great, but recently by making use of statistical methods, it may be said that the major parts of management are being carried out upon a comparatively solid basis. Thus, it can be seen that statistics has now become an indispensable means to a reasonable method for research work. Comparing this with statistics of former times, we cannot neglect the fact that statistics was considered as a material science, in other words, not as a formal science, and we dare say that this fact is a great change that has been wrought in statistics. If we allow for the fact that the late Mr Gosset[3] who anonymously supported R. A. Fisher's ideas,[4] we can not help being

[2] R. A. Fisher, *Statistical Methods for Research Workers*, 1. ed., 1925; 11. ed., 1950.
[3] W. S. Gosset, or "Student".
[4] R. A. Fisher, *ibid.*

surprised at finding how the principal features of statistics have so far been varied.[5] In the case of natural phenomena as well as social phenomena, it is impossible to determine the general nature (*Parameter*) of the facts at issue from an individual case; thereupon, we may collect comparatively many facts (*Samples*) which can be assumed belonging to the same category, and may seek for what nature they have in common, i.e., a typical one from the grouped phenomena.

Formerly, it was believed possible in natural science, e.g., through chemical experiments, to separate "water" of say one liter, and to ascertain its conditions and their interrelationships, but it was impossible to make such experiments in social science, because in studying social facts, being mass phenomena, it was only possible to make them clear by means of mass observation.[6]

But, nowadays, the research workers in physics, astronomy, biology, agriculture, medicine, etc., do not always content themselves with analysing only a few pieces of samples, but they have become accustomed to regard it as an ordinary research process to find out 'average' of the results collected by many observations. Darwin's evolutionism as well as Malthus's theory of population, may be considered as the analytical results of samples. As can be seen above, we may point it out, from the facts of progress in statistics, and of the abundance of statistical data, that the methods of modern science have undergone a change from deductive to inductive.

II

The facts above referred to, naturally relate to the fundamental principles of statistics or "Das Gesetz der grossen Zahlen." The nature of this law is to expose regularity, " 'Characteristic of population,' or 'Parameter'," which comes out clearly only in the case of great numbers derivable from the overall results of mass observation, i.e., it is to expose the regularilty of the relationships between their combinations in the static mass, and the regularity with which some events ought to occur, in the dynamic mass.

But in the case of the observation of small numbers drawn at random—mass observation consists of the results of these small numbers—the existence of such a regularity may not be recognized. Because statistics aims not at something individual but at something representative, i.e., "*Das Typishe*," and this goal can be fulfilled only by the observation of great numbers. Therefore, the more

[5] Mr. Gosset published his studies on sampling in the essay "*Probable Error of a Mean*" under his pseudonym (non de plume), "student." in 1908 Since then he continued his studies, and arrived at the correct solutions of the problems regarding distribution, besides 'Variance' i.e., the distribution of estimated value, the distribution of the average divided by the estimated standard deviation and the distribution of the estimated coefficient of the correlation between independent variables. With this essay, he broadly established the standpoint of the "χ^2" and "t" distributions in the theory of sampling. The development of the "z" distribution was a later result based upon these two. But in those days Gosset's propositions were by no means easily accepted generally, but after Fisher resolved them into simple arithmetic formulas through his incessant efforts, it may be said that it has become the common possession of mankind.

[6] Tyszka, *Statistik*, 1924. S. 27.

we increase the chances of observations, the more the deviations from the general average or the representative value decrease, in other words, the nearer can we approach its true phase.

Quetelet has explained the principle of the law of great numbers by a skillful example. Now, let us draw a distinct circle with a piece of chalk upon a large board, approach and watch it, then you will not fail to notice even a small spot. In other words, you may be able to distinguish every point which is different from one another somewhat incidentally or voluntarily, as great numbers. But the farther you step back from the board, the larger the points in numbers may be observed, and at the same time, you would be able to see that these points are divided into a certain length on the board upon which the circle is drawn, and if you step back farther, you are apt to overlook and have difficulty in recognizing the different combinations, and at last you may notice the general tendency which rules the variances of many points, i.e., the nature of a dotted line.

We may see that human community is almost identical to this. Though the human individual is allowed to be free within a certain limit, if we look at those voluntary acts from some distant place, we can hardly grasp them. Then we may easily recognize that only a general law is beginning to work.[7]

The law of great numbers was explained in another way by the Danish statistician, Westergaard, and the French mathematician, Laplace.

Assuming red and black balls have been previously put in a jar in the same number respectively, and every one of these balls in equal in size and shape, and is independent of each other. Moreover, they are by no means possible to be distinguished from one another even with the sense of touch. Suppose with your eyes shut, you take out a ball at random and see whether it is red or black, and again shut your eyes, place it into the jar, and after shuffling the jar many times, repeat the same process as many times. If this process is often repeated, then the probability that each of these balls is drawn, must equally be $\frac{1}{2}$ mathematically, because the same numbers of red and black balls were previously put in the jar.

Now, if such experiments are repeated 'S' times and if the red ball is drawn 'M' times and the black 'N' times, then $\frac{M}{S}$ or $\frac{N}{S}$ may tend to approach $\frac{1}{2}$ which is the mathematical probability (*a priori probability*), by increasing the number of 'S'. Such trials have been practised by students in many countries. The results, Westergaard obtained from his experiments, are as follows:

In the case of ten thousands drawings, the red balls drawn were 5,011 and the black 4,989, so that the frequency of drawing red and black balls is approximately the same in percentage.

Moreover, Westergaard divided the ten thousands trials into one hundred rows, which contained one hundred trials respectively, and he tested every of them in the frequency distributions. Then he obtained results that showed the

[7] Kaufmann, *Theorie und Methoden der Statistik*, 1913. S. 30; Tyszka, *a.a.O.*, S. 28.

lowest rate of red balls appearing was 34 and the highest rate was 63. Therefore, the rate was approximately 1 to 2. On the other hand, in the case of a row of five hundreds drawings the highest rate was 273 to 227, and in the case of a row of two thousands drawings, the highest rate was 1,041 to 959.

In presenting these as the rate of deviation, in the case of a row with 100 drawings, the extreme one was red balls, 34, and black, 66, and the absolute deviation to the average 50 being 16, its rate was 32%. On the contrary, in the case of a row with 2,000 drawings, the absolute deviation to the average 1,000 being 40, its rate was only 4%. Moreover, in the case of a row with 10,000 drawings, the abosolute deviation to the average 5,000 being 11, its rate was only 0.22%.

Thus, Mr Westargaard established a law similar to that formulated by Quetelet. In other words, it follows naturally that, in order to halve the divergence of uncertainty, the frequency of observation should be squared.[8] Quetelet, a Belgian, also made a trial like this, and Buffon, a French, counted the heads and tails of a coin tossed. By such a trial, he found that, out of 4,040 tossings, the heads appeared 2,048 and the tails 1,992, i.e., the frequency of the heads and tails appearing was 51% and 49% respectively.

Süssmilch asserted long before that, for the sake of determining the social

---

[8] Now, supposing the heights of 1,000 students, as a sample, are between 174.711 cm and 175.959cm, the uncertainty is the difference of these two figures, i.e., 1.248cm. In order to halve the uncertainty, 4,000 students in place of 2,000 should be studied. Again, if we want to make the uncertainty reduce to 0.5 cm i.e., we want to obtain the precision of $\frac{1.248}{0.5}$ or that of 2.50, we must study 2.50^2 or 6.25 times as many cases, i.e., 6,250 students. Therefore, the more the number of samples is increased, the more their confidence limits tend to decrease. And the confidence limits are inversely proportionable to the square root of 'N' in mathematics. Then we get the following:

$$\sigma_{\bar{x}}=\frac{\sigma_x}{\sqrt{N-1}}$$

This is the formula for the standard error of the mean. $\sigma_{\bar{x}}$ means the standard error of the arithmetic mean. σ_x means the standard deviation in population. $\sqrt{N-1}$ means the aggregate of the samples accounting for the degree of freedom. But if the statistical units in a sample are great enough, the difference between 'N—1' and 'N' may actually be neglected.

Thus, when the standard deviation is 6.58cm and the munber of samples is 1,000, then by using the correct formula, we get:

$$\sigma_{\bar{x}}=\frac{\sigma_x}{\sqrt{N-1}}=\frac{6.58}{\sqrt{999}}=0.208 \text{ cm}$$

In the case of normal distribution, the areas enclosed by the base-line under the curve, show the following relations: The area enclosed between the mean and ± 1 standard deviation is 0.6827 The area enclosed between the mean and ± 2 standard deviations is 0.9545. The area enclosed between the mean and ± 3 setandard deviations is 0.9973.

Thus, the area enclosed between ± 1 standard deviation is approximately two-thirds of the total area, and the area between ± 2 standard deviations is about 95 percent of the total area, and the area between ± 3 standard deviations occupies almost all of the area. Thus, when the distribution of sample means shows the normal distribution, and the mean of sample means, '*statistics*' coincides with the accurate mean of population, '*Parameter*,' the following relations will be found from the nature of normal distribution, i.e., the chance in which the error of statistic is greater than one standard deviation from parameter, i.e., one standard error, occurs once in three, and the chance in which the error of statistic is greater than two standard deviations from parameter, i.e., two standard errors occurs once in twenty, and the chance in which the error of statistic is greater than three standard errors scarcely occurs. In other words, the highest limits of error may be said to be three standard errors.

law, great numbers should be observed, and he said as follows:

"In a small circle, everything appears to exist without any order. In order to show up the hidden law, one should observe as many cases as possible extended to the whole circle.[9] Süssmilch believed that, not only the greater the mass, but also the longer the time, the clearer the constant cause tends to become apparent, and thus, a certain law or constancy in social phenomena clearly emerges. And, Süssmilch recognized the existence of such a law, and attributed it to the order of God, so that we may conclude that he related the law of great numbers to his faith. From the point of view of our physical life, there are outstanding divergences arising from the various causes, in the number of the births of males and females among families. But, based upon mass observation, the divergences of the birth-ratio between the sexes tend naturally to become diminished. If the mumber of observations, and the period of investigation are great enough, the birth-ratio between the sexes tends almost to be constant. For example, in Japan, the birth-ratio between the sexes ranges from 102 or 103 of male births to 100 of female births, and thus we may say that the birth-ratio of males always exceeds that of females. Moreover, such a ratio may come out not merely in Japan but also in any part of Europe, and thus, we can recognize that it is a general feature. But as the death-rate of the male in his childhood is generally high, compared with that of the female, the ratio between the sexes is approximately able to maintain its balance.

In the case of death, there is also a certain regularity or constancy. And there is a certain number of deaths in each age, the deviation from that number each year is considerably small; in addition, these relations will continue to exist for a long time to come, a decade or a century.

Thus, we may mathematically be able to calculate the death order of life, the average length of life, average duration of life, etc.

Unless we rest on the facts referred to above, we can not recognize our constant order of death, which, in practice, has a very important meaning in life insurance. We have so far seen constancy in our physical world, but there is also constancy in our voluntary activities and they are similar to each other. These points were already referred to clearly in Wagner's well-known work, "The law in the Apparently Voluntary Human Activities",[10] which is worth reading even today.

There are also certain rules in marriage, divorce, suicide, crime and the like. In marriage, the will and the wish of each person may be of a quite different sort, but nevertheless, the total number of marriages in any country differes hardly from year to year. In divorce, it is similar to them in marriage. If, by any chance there is a variation, it will be very gradual, it rarely changes by incident, or suddenly. It is also the same with suicide, i.e., the number of cases of suicide amongst the sexes maintain an approximately constant ratio in any year, and its fluctuation

[9] J. P. Süssmilch, *Von der Göttliche Ordnung*, S. 64; Tyszka, *a. a. O.*, S. 28.
[10] Adolf Wagner, *Die Gesetzmässigkeit in den scheinbar wilkürlichen menschlichen Handlungen vom Standpunkte der Statistik*, 1864.

is very small. The more enlightened the people become, the more the cases of suicides tend to increase, but the tendency is gradual, and the trend is slow and regular, not in an extraordinary or unexpectable way.

It is so with crimes and illegal acts; in other words, there are like numbers repeated in the number of plaintiffs and defendants, in law suits, and their judgment ratio are repeated in a similar way.

Quetelet is quite right even today when he said, "there can not be two budgets, so we must pay regularly for equipment every year, such as prisons, guillotines and galleys."

There is nothing different from that in economic life too. Here we may find regularity and constancy everywhere in our economic life. Though there are in fact certain changes in the production and the consumption of goods, these changes are extremely gradual as well as permanent. With the public budget and the cost of living, the revenues and payments for this year rest mainly upon those of the past, and these processes are to be repeated. Thus, both in social and in economic life, we may see everywhere a certain regularity and constancy.

Needless to say, in these cases, it must be submitted that the number of observations is great enough.

Bortkiewicz asserted a parallel law or law of small numbers, in contrast to the law of great numbers. Thus, constancy or stability will arise from its own nature in a small series, as well as in great numbers. Child suicides or accidents with casualities will show relevant examples. He firmly asserted that in these cases, their real regularities will approach the mathematical probability, and his assertion may be said to fit the reality.

But the so-called law of small numbers is nothing but a specific case of the law of great numbers. Because in this case, the fact that real regularities will approach the mathematical probabilities, can only be certified when small series connect with great numbers. In other words, when investigation is carried out upon a large area, constancy and similarity will occur even in small numbers.

The so-called 'similarity in the world' has an important significance from the practical view-point as well as from the scientific view-point, and therefore, it refers naturally to the principles of probability.

III

It is the problem of probability, "*Wahrscheinlichkeit*," which is closely related to the law of great numbers. The problem of probability above refereed to is nothing but the theory of probability in mathematics. On this point, Laplace, a Frenchman said, in his treatise[11] that when the number of observations of so-called 'chance'—the cause why chance happens of which we are out of the reach

[11] Laplace, *Essai philosophique sur la probabilité*, 1814.

of exploration—is increased, a remarkable regularity might tend to occur. Laplace said in a passage of his essay as follows:

"Almost all of events which seem to be strange to the natural law, constitute a series of natural or necessary phenomena, just as the sun constitutes the center of the universe. As we have no knowlege about the relationship of the whole world, we might rather attribute the same events to fitness, 'Zweckmässigkeit' or chance. But there are no "chances" at all in the world. To attribute them to chance is nothing but to explain that we are ignorant. World phenomena are ruled not by the outside intervention of a spirit but by the sole law of causation."

This fact that phenomena are of regular frequency—we call them "chance" because it is impossible to know the causes—makes for us mathematical dealings possible, in our actual life. It is the so-called theory of probability which we want to apply.

Thus, the theory of probability is one of the principles in mathematics, and is the most excellent method as an instrument for verification within the scope in which we can deal, by ourselves, with those problems to solve, as we have no complete knowlege.

Thus, the mathematical probability stands against hypothetical judgment or logical necessity. With the latter, when 'a' exists, there must be the existence of 'b'. For example, supposing there is a triangle, the total of three inside corners must be equal to two right angles.

But with probability, disjunctive judgments are accompanied, i.e., if there is 'a', there must be any of 'b', 'c', 'd', 'e' or 'n'. For example, in the case of casting a die, any one of the spots, 1, 2, 3, 4, 5 or 6 will appear. In the case of birth, there is either a still-birth or normal-birth, of either male or female. The necessity of such a case is connected with not only the occurrence of one event, but also the occurrence of any of these events. Moreover, one of these events must occur, therefore, the task of calculating probability is to ascertain the precision of the frequency of probability. Basing ourselves upon experiencial facts, lacking comprehensive causes, (*Parameter*), the statistical averge is as in the case of *a priori* probability, sought to draw out a certain rule or regularity from a phenomenon.

The so-called 'similarity in the world' rests on a statistical average based upon the theory of probability. But a statistical average differs to a certain extent from what is calculated only by the mathematical method. In other words, the theory of probability in mathematics deals with almost all of the factors in the world.

With the statistical average, as there are some limits, it can not generally deal with all of them, from the statistical view-point; this is called 'statistic' of sample deviation.

As a result, not only in physical or natural phenomena, but also in social phenomena, departing from the theory of probability, it may sometimes occur that the males' birth exceeds the femalers' or vice versa.

But, generally it may not be denied that statistical averages play an important rôle in social or economic life. For example, in our economic life of today, the business man relies generally on the results obtained from the statistical data of production, and acts with his expectation that there will neither be a social fluctuation nor will there be a very little shift, if any, in the forthcoming years.

If we can not have such an expectation in our social or economic life, such activities will come to a standstill and would be disturbed. Here we can see that the theory of probability is closely connected with statistics.

IV

Regarding the fact that there is a rule or constancy in natural as well as social phenomena, it has not so far been dealt with in England, although this fact was a big problem for a long time in the Continent, especially in Germany. Thus, the author of 'Göttliche Ordnung,' Süssmilch said that all phenomena in the Universe are ruled by the God's order, which He predeterminedly revealed, so that not only the balance of the sexes are maintained, but also population, which is gradually increasing from year to year, and the food required to maintain the population could be balanced.

Altxander von Öttingen made a trial to solve this problem in his "Die Moralstatistik" from the view-point of his teleology and religion in part.

On the contrary, Quetelet and his disciples intended to solve this problem from the view-point of natural science. In other words, they thought that the law prevailing in social phenomena was the natural result accruing from the law of nature. Thereupon, they concluded that self-determination was merely superficial and the truth lay in our predetermined destiny. This fact seemed to be justified not only in the process of physical life, such as birth and death, but also in almost all of the activities which were believed to be the result of our free or voluntary judgments.

Prof. Adolf Wagner, at the Berlin University, who was well known as a modern economist, also denied the idea of 'Free Will' in his essay above referred to. To quate a passage: "But then, it is noteworthy that, though we work as the faithful members of a great mechanism, we have yet a restricted free movement, without preventing such a mechanism from moving, moreover, we believe that we are doing perfectly self-determined activities beyond such a limit. But generally from the viewpoint of mass, our activites are ruled by rigid and general causes and occur just in the same way as the physical world-order moves."

Now, if there exists really any paradox between these two, can statistical law deny the existence of 'Free Will'? Naturally, there is no paradox between them, and if we observe them in detail, this problem is quite easy to be solved, so that we need not seek a way to supernatural religion, or metaphysical reason or fatalism because at first the statistical average has no relation to 'Free Will,' and secondly, 'Free Will' is a problem of ethics and we have 'Free Will' from the view-point of morality. To illustrate in detail, we act according to our own judgments and opinions, and such an act is called 'Free Will' from the ethical point of view.

Man's free acts arise from the commands of his own reason, thereupon man is inclined to behave with higher and nobler motives.

On the contrary, an act ruled by the animal wants is called "Unfrei." There is always a motive in any of our activities, but in almost all cases, the motive is not simple, but is so remarkably complicated, that it is quite useless to try to find out the cause of each activity.

Therefore, in examining statistical regularity, we should make it our object to observe the whole or mass phenomena instead of individual acts. Only through such a method of observation, can we successfully exclude any and all individual activities, arising from chance, and we can get the general rule coming out of the constant causes which are free from the peculiarities due to space and time. Then our 'Free Will' is kept balanced, but never annulled.

Prof. Schmoller explained this as follows: "To-be-free, means that we ardently push to the front to be a man of ethical being. If it is true, the superior free means not voluntariness but the norm. The norm is said to be the absolute good and ideal, then we will call the act with the strongest motive, *the most free.*"

Reference Books:——

1. Carl von Tyszka, *Statistik*. Jena 1924. S. 26–39.
2. Arthur L. Bowley, *Elements of Statistics*. 6.ed. London, 1948. Chap. 3, pp. 3–17, 287–311.
3. Albert E. Waugh, *Elements of Statistical Method*. 3. ed. New York, 1952. pp. 2–6, 155–185.
4. Otto Donnor, *Statistik*. Hamburg 1937. S. 11–16.
5. R. A. Fisher, *Statistical Methods for Research Workers*. London, 1950. pp. 1–23.
6. Frederick C. Mills, *Introduction to Statistics*. New York, 1956; Chap. 6; Chap. 7. pp. 134–191.
7. Wm. L. Crum, A. C. Patton and A. R. Tebutt, *Introduction to Economic Statistics*. New York, 1938. Chap. 13, pp. 208–220.
8. Yuzo Morita, *General Theory of Statistics* (in Japanese). Nihon-Hyoronsha, Tokyo, 1941.
9. Takuma Terao, *An Outline of Statistics* (in Japanese). 2. ed. Keio-Publishing Co., Tokyo, 1948.