

**LADISLAUS VON BORTKIEWICZ,  
PROBABILITY AND STATISTICAL  
STUDIES ACCORDING TO KEYNES\***

ŚLĄSKI  
PRZEGLĄD  
STATYSTYCZNY  
Nr 17(23)

Oscar Sheynin

International Statistical Institute

e-mail: oscar.sheynin@googlemail.com

ORCID: 0000-0002-5039-4314

ISSN 1644-6739  
e-ISSN 2449-9765

DOI: 10.15611/sps.2019.17.06

*Wahrscheinlichkeit und statistische Forschung nach Keynes,*  
Nordisk Statistisk Tidskrift, Bd. 2, 1923, pp. 1-23.

**Abstract:** Keynes was a versatile scholar and his reviewed work is widely known. Bortkiewicz notes that it was mainly concerned with the logical foundation of probability and methodology of statistical judgement. Keynes paid significant attention to the principle of indifference, unmeasurable probabilities, and to contrasting calculations with circumstances. In his paper he reproached Bortkiewicz for the estimation of precision in statistics (Bortkiewicz did not comment) and criticized both his manner of writing and his law of small numbers.

**Keywords:** law of small numbers, logical foundation of probability, principle of indifference, unmeasurable probabilities.

Keynes' work [1921] was reprinted at least twice, in 1952 and in 1973, as volume 8 of his *Collected Writings* (there also exists an eBook of 2014, No. 32625). The paging in these sources naturally differs. Bortkiewicz certainly referred to the first edition but this author has only seen the other two editions and was able to check those references in a few cases only.

All thinking statisticians, whether working near or far from mathematics, have every reason to wish to ascertain for themselves the relation of the statistical method with the notions of probability and probability theory in general. It is also necessary for them to study the very different opinions about these relations. Some tend to consider all scientific statistics as the applied theory of probability, whereas others think that that theory should only be applied in special problems of statistical studies. This is a problem which will be discussed for a long time. Keynes also skilfully considered it, although his book was not wholly devoted to that problem. Only its last, fifth part on the theory of

---

\*Based on [Keynes 1921].

statistical conclusions is directly concerned with statistics. The first four parts discuss, respectively, opinions based on probability; the logical basis of the theory of probability; induction and inferences by analogy; and *some philosophical applications of probabilities*.

There exists however a close connection between Keynes general considerations and his special thoughts about the statistical method, so that his book is an integral contribution, quite original both in its intention and execution.

In his publication Keynes does not conceal at all the origins of the necessarily borrowed building materials, he indicates them scrupulously. He mostly feels himself akin in spirit to his English predecessors. Like them, he attempts to avoid the areas of creative fantasy and to keep himself always connected with the factual material. He therefore rejects the Laplacean enthusiasm<sup>1</sup> to which his contemporaries had gladly listened and for a long time remained in his thrall.

Ellis [1842] was the first who objected to that *alchemy of logic*. Then, Venn [1866] developed a theory wholly based on empiricism. Actually, probability appeared there only as the statistical frequency of some event and, moreover, only when it is precisely determined by observations continued infinitely.

Keynes however resolutely objected to that *frequency theory*. It was the *empirical school* that went too far in its response to the viewpoint of Laplace. Were our experience and science perfect, the theory of probability would not be needed at all.<sup>2</sup> However, if gaps are discovered in our knowledge, then, as Keynes supposes, probability-theoretic judgements become inadmissible if not supported *either* by this line as incomplete intuition *or* by something else, i.e. by a specially intended

---

<sup>1</sup> Laplace had indeed enthusiastically described his humanistic political views partly based on general stochastic ideas and on the compensating, as this author would say, action of randomness in mass random phenomena. We do not know what exactly Bortkiewicz had in mind, but in any case he (like many other commentators) had here and below somehow belittled Laplace whose initial views were allegedly reduced to the introduction (after De Moivre and actually even after Jakob Bernoulli) of the *classical* definition of probability. First, however, until the advent of axiomatics that definition remained in use, although justly rejected, and became only rivalled by Mises.

Second, Laplace had not at all restricted his activities to introducing that definition, he repeated many times over that hypotheses (perhaps including the number of cases as well) ought to be incessantly corrected by new observations, see for example his *Essai* [1814/1995, p. 116]. This does not correspond with his statement that probability rests on incomplete knowledge.

Finally, Laplace [1812, Chapter 5] considered geometrical probability and (ibidem, Chapter 6) solved some problems by the Bayesian approach.

<sup>2</sup> However, mass random phenomena can only be studied by the theory of probability. Then, instability of motion, and especially the newly studied chaotic movement, prove that that statement was wrong.

prior principle. The author will show in this article that in his own doctrine Keynes admitted both as the principle for basing probabilistic opinions.

In spite of his strong and pointed opposition to Laplace, Keynes believed it necessary to join him here. According to Laplace, each numerically measureable probability (and he supposed that probability is such, indeed) rests on an incomplete knowledge of circumstances which makes it possible to distinguish two or more mutually exclusive cases whose occurrence is equally indefinite. This is achieved by calculating the probability of an event as the ratio of the number of cases favourable for it and the total number of all the possible cases. This principle of defining probability is referred to in different ways. Boole [1862/1952, p. 390], for example, mentioned an *equal distribution of our knowledge or rather of our ignorance*.

Von Kries [1886] chose the principle of *deficient grounds*, but Keynes considered this expression unsatisfactory and suggested *principle of indifference* (p. 44 of the eBook). Actually, Keynes clearly agreed with the long-standing criticisms of this principle by von Kries although only insofar as the Laplace formula still held. Indeed, he believed that the principle of indifference should only be introduced more rigorously after which it would become a suitable and single useful foundation of probability theory.

1. The Keynes viewpoint can be satisfied if we say that he contrasts the *simple* and the *special* forms of the principle of indifference. However, he offers some reservations. The most essential of them discusses irrelevance. Thus, for proposition  $x$  which is based on result  $h_1$  the circumstance  $h_2$ , if the probability of  $x$  (more precisely, the probability that  $x$  is corroborated) does not change when  $h_2$  is added to  $h_1$ . Or, more generally, if something follows from  $h_2h_1$  (i.e., from  $h_1$  when  $h_2$  is added to  $h_1$ ), but not from  $h_1$  alone. On the contrary, in such a circumstance  $h_2$  is relevant if it alone, or some corollary of  $h_2h_1$  somehow changes the studied probability.

We ought to note that Keynes considers  $x$  and  $y$  as well as  $h_1$  and  $h_2$  as propositions and therefore sometimes applies symbol  $f(x)$  for  $h_2$  which denotes a proposition connected with  $x$ . For justifying an equal probability of two different propositions  $x$  and  $y$  by a given result, this result should not include any such circumstances which have to do with  $x$  or  $y$  if they do not correspond to a suitable circumstance *of the same form* bearing on  $x$  or  $y$ .

In a symbolic form this means: if  $f(x)$  is included in  $h_1$ ,  $f(y)$  should also be included there and vice versa<sup>3</sup>. Keynes explains this rule, which

<sup>3</sup> Bortkiewicz unsystematically connected  $f(x)$  with both  $h_1$  and  $h_2$ .

for the sake of brevity can be called the *rule of symmetry*, by an urn with a certain number (say, 4) of black and white balls. It is asked (and Keynes offers both solutions) should we consider five numbers, 0, 1/4, 1/2, 3/4, and 1, which are the possible ratios of the black balls to their whole number, equally probable, or believe that each ball can equally be black or white so that those ratios will be 1/16, 1/4, 3/8, 1/4, and 1/16 (...). Keynes decides in favour of the second solution but his statement is based on a vicious circle.

Keynes indicated that, in particular, von Kries and Stumpf had studied that example<sup>4</sup>. The former decided correctly, in this author's opinion, that that example resolutely does not at all admit any numerical determination of probabilities. (...) Stumpf's criticism induced von Kries [1916] to return to this example and he attempted to strengthen his former viewpoint by new considerations.

Such thoughts only confirm what was clear from the very beginning: it is futile to attempt here a derivation of a general exact numerical solution, and since Keynes busies himself with such a problem, we may ask ourselves: does this attempt correspond to the need to be supported by facts (by *matter of fact*), to the principle which he adopted?

It is also surprising that in this special case Keynes sided with the *subjectivist* Stumpf rather than with the *objectivist* von Kries<sup>4</sup>. True, in general he is much closer to the latter, but incidentally, he expressed some dissatisfaction with von Kries' discussion of the logical foundation of probability theory he himself highly estimated.

This possibly happened at least partly since he only considers the book of von Kries [1886] but not a series of his papers [1888] or his *Logik* [1916]. May we wonder that Keynes was barely successful with the main, as he thinks, condition of the principle of indifference when studying the example in which the fruitfulness of that condition was necessary to reveal? This author does not think so. Does not the symbolism which Keynes applies with such a special liking imitate the non-existing precision?

The notation  $f(x)$  and  $f(y)$  ought to show that the forms of the pertinent propositions connected with  $x$  and  $y$  respectively are the same. But how should we understand the *form*? In the example now being considered, Keynes thought that those forms were different because one of them stated something about the number of combinations out of four taken two at a time whereas the second, out of one at a time. Does this not testify to the surprisingly petty interpretation of the expression *form*?

---

<sup>4</sup> Keynes refers once to von Kries and three times to Stumpf, see Bibliography. The author added two other contributions of Kries since they are also needed. Bortkiewicz [1899/17] publicly criticized Stumpf, too.

The rule of symmetry which Keynes had established becomes no more definite, it only allows us to see whether the circumstance  $h_2$  is relevant or not (p. 111). Intuition ought to support the application of that rule (pp. 53-54 and 64); in addition the result  $h_1$  invariably includes an inexhaustible set of various circumstances so that here only intuition can help us once more.

A similar difficulty certainly appears in each inductive study of the causes of phenomena but the practical inexhaustibility of  $h_1$  remains as a circumstance which is still not much facilitated by the rule of symmetry. According to Keynes this rule is the most important, but nevertheless not the only means for attaching more rigour to the principle of indifference.

Other measures of precaution also exist, but they do not change the fact that under some circumstances that principle can lead to differing, but still to some extent equally justified, probabilistic decisions. Keynes himself does not maintain that it is possible to confirm that principle by a unique solution [elimination] in *all* cases of such a contradiction, a circumstance which critics had expressed against it. It is already somewhat surprising that Keynes (p. 52) so earnestly attempts to *rehabilitate* that principle.

3. At the same time we ought to take into consideration that in the Keynes theory it is far from occupying the governing place as it did in the Laplacean theory. The latter justified the very notion of probability on the concept of *equal indecision*. In Keynes' terminology, on the principle of indifference which is most closely connected with Laplace's understanding, on the principle that each probability is a proper fraction. The structure of equally possible cases which serve as parts of such fractions is indeed based on the principle of indifference.<sup>5</sup>

On the contrary, Keynes' starts from a much more general notion of probability and his subject is thus not the theory of probability as such (numerous formulas in his book mostly concern not its field but logical calculations), and not only its logical foundation but the general methodology of probabilistic judgement. The doctrine of probabilities in the Keynesian sense (p. 97) is concerned with logical causes which prompt us to trust something rather than anything else. He (p. 98) bears in mind all kinds of arguments which issue from some assumptions and lead to reasonable although not certain conclusions.

---

<sup>5</sup> Laplace [1812/1886, p. 181] adds a reservation: probability is that ratio when there is nothing for believing that one case appears more often than the others. Then he [1814/1995, p. 6], however, states that the definition of probability is a principle [and does not therefore need any justification].

In his work, Keynes distinguishes measurable and immeasurable probabilities<sup>6</sup> and considers both. He believes that we are often incapable of comparing the degrees (Grad) of the latter, and wittily justifies this statement (pp. 34-40). If we agree that not each probability can be numerically expressed and if, on the strength of the principle of indifference, the probabilities of some conclusions obtain unequal numerical values depending on the approach to the appropriate problem – then we miss that strangeness which in such cases is mostly inherent in them. Contrasting solutions are sometimes based on different premises so that the values of their probabilities cannot be united by a single, so to say, higher value of the probability. Indeed, it is required here that the probabilities of the corresponding premises be numerically represented, which is not occurring here.

Keynes is strongly convinced that the immeasurable probabilities exist along with the measurable ones, and that they have precisely more weight in each kind of inductive studies, and this viewpoint is revealed in his entire discourse, both when he develops his own thoughts or criticizes alien viewpoints. This belief indeed predominates in the last part of his book. Each time when a statistical *characteristic*, be it a relative number, a mean value or a correlation coefficient, is transferred on unobserved cases (which is really the *inductive*, as opposed to the *descriptive* role of statistics) this method leads to a result which lacks any certainty and only possesses a better or worse justified probability.

4. It is this probability that cannot be expressed in any quantitative form and Keynes (p. 367) who specifically refers to Laplace, mentioned the attempts to allegedly specify the probability of some inductive corollary by formulas of probability theory:

*We will (...) endeavour to discredit the mathematical charlatanry by which, for a hundred years past, the basis of theoretical statistics has been greatly undermined.*

Keynes (p. 369) quoted Leibniz:<sup>7</sup>

*Estimation of probabilities is extremely useful, although in examples taken from the law or political sciences' delicate calculations are not as necessary as an exact listing of all the circumstances.*

Keynes also noted that the essential in his views was thus expressed by Leibniz. He had thus largely agreed that in statistics a quantitative application of certain formulas of the theory of probability often led to

<sup>6</sup> Bortkiewicz should have referred to Cournot [1843, §§ 233, 236] who had introduced unmeasurable (*philosophical*) probabilities. It is hardly known that Fries [1842, p. 188] forestalled Cournot.

<sup>7</sup> The author has translated this Latin piece from its German translation [Gini 1946, p. 405].

serious abuse, but we still ought to ask, had not Leibniz gone too far in belittling calculations in general and did not Keynes too eagerly agree with him here. The quotation from Leibniz is taken from the supplement of 3 December to his letter of 26 November 1703 to Jakob Bernoulli, in answer to the latter's letter. There, Bernoulli, in particular explained by example the method of empirically determining probability (...).<sup>8</sup>

Were these arguments of Bernoulli not decisive? He was certainly tending to develop the theory of probability as seen in his thesis [1685/1969, pp. 269-270]:

*Quanto caeteris scientiis praestet [he discusses mathematics] vel ex eo constat, quod cum reliquae de rebus, in se certissimis ac constantantissimis, non nisi probabiliter, illa de rebus maxime fortuitis et casualibus, v. gr. sortitionibus, apodictice et certissimo ratiocinio discurrit.*

Even if we disregard such extreme statements and remain within the boundaries of the dispute between Bernoulli and Leibniz, it is still impossible, as the author thinks, not to agree with the former whereas the latter does not look so well.<sup>9</sup> Once more he revealed his poor understanding of the theory of probability which surprisingly contrasts with his predilection for combinatorics (noted previously by Couturat<sup>10</sup> et al). Indeed, he thought that 11 and 12 points in a throw of two dice were equally probable! He certainly was unable to solve more difficult problems, as for example the calculation of the present value of an annuity [Bortkiewicz 1907/44, pp. 71-72n].

In his correspondence with Jakob Bernoulli he unsuccessfully contrasted *calculations* and *enumeration of circumstances*. Apart from games of chance the sought probabilities cannot be derived from circumstances.<sup>11</sup> Their most accurate regard can therefore be only understood when it is intended to go over from probabilities derived from observations to unobserved cases while taking care of ensuring the most possible coincidence of the general conditions. This is troublesome.

But the observance of these directions does not change anything in the essence of the necessary calculations and will mostly only lead to their justification by observations restricted to a tighter area, and correspondingly to dealing with smaller numbers.

<sup>8</sup> Bortkiewicz describes Jakob Bernoulli's reasoning in great detail although a German translation of the *Ars Conjectandi* had already appeared (in 1899).

<sup>9</sup> Calculations should not be opposed to the study of circumstances. This author notes that Gauss [*Werke*, Bd. 12, pp. 201-204] stated that the nature of the studied object ought to be taken into account. W. E. Weber described his opinion in a letter of 1841 [Gauss, *Werke*, Bd. 12, pp. 201-204].

<sup>10</sup> Possibly Couturat (1901). Slightly further Bortkiewicz mentions a surprising mistake made by Leibniz, see Todhunter [1865, p. 48].

<sup>11</sup> The sought probabilities do not always exist.

It is perhaps somewhat interesting that Keynes (p. 268n) strongly and possibly too strongly criticized Mill [1843/1886, p. 353], censured him for the method (surprisingly similar to Leibniz's way of thought) of taking circumstances into account:

*Even when the probabilities are derived from observation and experiment, a very slight improvement of the data by better observations or by taking into fuller consideration the special circumstances of the case, is of more use than the most elaborate application of the calculus of probabilities founded on the data in their previous state of inferiority. The neglect of this obvious reflection has given rise to misapplications of the calculus of probabilities which have made it the real opprobrium of mathematics* (Mill).

Certainly so, since mathematicians repeatedly applied statistical numbers without more thoroughly checking their reliability and in addition often unjustifiably and too sketchily dealt with the peculiarity of the objects of study.

In this sense Mill's statement is indeed worthy of attention but it concerns the basis, i.e. the legitimacy, the structure and formulas of probability theory in their application to statistical materials as little as the remarks addressed long ago by Leibniz to Jakob Bernoulli.

5. Keynes himself says however that *Leibniz's reply goes to the root of the difficulty*. Full of anti-calculation tendency, which he thus possesses along not only with Leibniz, but with Mill, Keynes particularly reproaches the estimation of precision in statistics, i.e. those calculations which ought to establish tentative or limiting deviations of the obtained results from reality.<sup>12</sup>

Czuber [1910, pp. 15-16], for example, resulting from the number of boys and girls born during a certain period calculated the most probable number of new-born girls for a later period in Austria after assuming that the number of boys was known. Applying some formulas of probability theory he established a very high probability, barely distinguishable from certainty, that the real number of new-born girls will be comparatively very near to the number provided by him.

Keynes first of all criticized Czuber for transferring the sex ratio onto a larger number of births and for believing that in spite of this circumstance, the result was practically true. This, as he stated, contradicts *good sense* and some theoretically derived requirement.

Second, without sufficient grounds, Czuber thought that this sex ratio was stable. To what extent this is inadmissible is already visible from the

<sup>12</sup> Bortkiewicz (1894-1896) had earlier denied the need to estimate precision [Sheynin 2017, p. 176]. This theme is treated in detail below.



later period of 1895-1905 [Keynes, pp. 351-353], deviations were situated outside the limits to which the same method had attributed practical certainty.

We only ought to discuss this second objection since it is much more important. Indeed, the reference to *good sense* in the first one was unconvincing, whereas the discussion of the *requirement* would have led us too far.

In essence, Czuber's calculation assumes that the sex ratio has a normal dispersion, that is, possesses the highest possible degree of stability. For this reason his results are to a certain extent usual and he himself (p. 13) makes it known when he preliminarily refers to the

*previously generally assumed notion about the constancy of statistical relative numbers.*

Hence on p. 16 he mentions the subsequent modifying statement. Keynes had not taken the context into consideration. Czuber himself was not touched upon by Keynes, but in its essential part the criticism remains completely valid in respect to a countless number of calculations if a direct practical meaning is attached to them. Nevertheless Keynes had not said anything new.

Already Venn reasoned about the transfer of statistical frequencies from observed to unobserved cases, and especially about transferring them to the future, when the assumption of the constancy or of only insignificant change of the general conditions is decisive.<sup>13</sup> This is the cause of the uncertainty which, as he formulates it, belongs to the area of induction rather than probability. This should mean that a mathematical approach by the rules of probability theory to uncertainty caused by the assumption of stability is useless, see Mill [1843, Chapters 1 and 6].

It is somewhat surprising that Keynes, who thoroughly discusses Venn and considers him to some extent as a precursor of Lexis (see below), does not say a single word about his important reasoning. Incidentally, it appears to this author that in general Keynes describes Venn's views somewhat crudely. In essence he pays no attention to the (very important for Venn's viewpoint) mental replacement of actual by *imaginary series*. The *frequency theory* thus loses its excessively empirical side, although naturally without establishing its certain basis. In

---

<sup>13</sup> About 1965-1968 the author was in touch with Login Nikolaevich Bol'shev who was Kolmogorov's student and who around 1967 became head of the section on mathematical statistics of the Academic Mathematical Institute. He once said that because of financial considerations it was even necessary to make conclusions after only one observation. It was the author's impression that he spoke about some military matters. Regrettably, Bol'shev died prematurely.

order not to be considered a follower of that theory, the author's intention is indeed to note that circumstance.

That theory suffers, in the author's opinion, exactly because of the attempts to reduce the notion of probability to two ideas, to an *irregular order* of the elements of a series and to the *approach* of the appropriate empirical frequency to a fixed boundary value which can only be more precisely indicated by probability theory. Neither Keynes nor Bosanquet, a long-standing opponent of the frequency theory, had noticed its indicated weaknesses.<sup>14</sup> We may consider Venn its main representative and therefore the author is in complete agreement with Keynes about its conclusion, i.e. agrees to deny it, if not quite in regard to its causes.

Nevertheless, Keynes' criticism of Venn seems as though one Cambridge man is not quite impartial to another one. As Keynes demonstrates to him, Venn is hardly characterised by a *remarkable acumen* [Edgeworth 1911, p. 403; 1996, vol. 1, p. 152].

6. Some uncertainty unyielding to calculation is inherent in each transfer of the statistical frequency from observed to unobserved cases. Lexis later confirmed this idea by studying the real behaviour of series of population and moral statistics. This idea was precisely formulated in the theory which he had derived for explaining his results. In accordance with that theory, apart from rare exceptions, the indicated kind of uncertainty is caused by the physical or essential component of fluctuations which acts along with its normal random or non-essential component.

Only the first, but not the second component, as Lexis reasoned, admits a probability-theoretic interpretation. It immediately follows that in statistics, estimations of precision, insofar as they inevitably only consider the non-essential and miss the essential component, are therefore illusory.

---

<sup>14</sup> Bortkiewicz mentioned the English philosopher Bernard Bosanquet (1848-1923). The true originator of the frequentist theory was Mises. Here are his words [1928/1972, pp. 26-27]: "His theory is not *quite new*, Venn had *thoroughly described* the idea of determining probability by frequency. His other forerunners were, as he stated, Fechner and Bruns who had proposed a doctrine of the collective. However, the *clearest* in this direction was Helm (1902), but still neither these authors nor many others were able to create a perfect (?) theory of probability since they had not introduced disorder, the *decisive* indication of the collective." Note however that in the introduction to the 1931 edition of his book Mises named Ellis and Cournot as his predecessors. Mises regrettably omitted Bayes [Sheynin 2017, pp. 68-69].

A little earlier Bortkiewicz mentioned disorder as a weakness of the frequency theory which only makes sense only insofar as a mathematical definition of disorder was not offered even now. Mises very easily surmounted the second weakness, as Bortkiewicz called it, by defining probability as the limit of frequency.

Already Venn declared that *the causes of uncertainty in the area of induction* are the more significant the longer the series of unknown cases for which a certain frequency is postulated. This statement should have meant that in statistics the estimation of precision becomes the less reliable, the larger the number of the appropriate observations.<sup>15</sup> The same conclusion follows from the Lexian theory, according to which the essential component ever more violently perturbs the fluctuations as that number increases.

However, Venn thinks that that influence is caused by the lengthening of the period of observations with the increase in that number, but Lexis explains this phenomenon more generally: the field of observation can be extended by a longer period of observation, by widening it, by a larger number of large groups (?) and each such possibility leads to a stronger action of the essential component of fluctuation.

We may therefore say that in general an estimation of precision is the less suitable the more observations serve as its basis. We would also refer to the reasoning of von Kries [1886, pp. 178-181] and Karl Wagner [1898]. Keynes had not included the latter's contribution in his bibliography, but mistakenly attributed a paper of the same author to Adolph Wagner, a specialist in life insurance.

Keynes does not at all indicate that in statistics, the significance of the estimation of precision exactly in the sense which follows from the Lexian theory of dispersion depends on the number of observations. Had he drawn on that theory for judging the suitability of the estimation of precision, he in any case would have been more conciliatory inclined and possibly acknowledged that in the struggle against such estimates Leibniz does not justify hopes as an ally.

Indeed, as stated above, the Leibniz postulate leads to the most precise registration of the *circumstances* so as to work with a relatively smaller number of observations.<sup>16</sup> However, the smaller that number, the

---

<sup>15</sup> The statement in the last sentence is repeated below. A comparison with the treatment of observations is possible: random and systematic errors are as though two components of fluctuations, but the estimation of precision in the case of a small number of observations is difficult.

Now, unobserved cases became observed. Then, the Lexian theory allegedly leads to the increase of the physical component of fluctuations with the number of observations (below), but Lexis made no such statement. On the contrary, he [1879, § 15] qualitatively asserted that the more observations there are in each series, the nearer  $Q$  is to unity and the smaller that component, see also his formula in § 11. Elsewhere Bortkiewicz [1931, p. 5] stated that the physical component tends to increase with a decreasing number of observations.

<sup>16</sup> This assumption is too restrictive. Concerning the next sentence cf. the beginning of Note 16 above.

more admissible and even the more indicated the estimation of precision, and the more we are induced to apply the formulas of the theory of probability! It is not accidental that such a specialist in similar investigations as Westergaard is among those who speak out for estimating precision and often apply it.

7. However, if Keynes, as the author suspects, unjustly refuses to adjoin Lexis this certainly does not happen because of his objection to the Lexian direction. He is rather earnestly interested in Lexis, stresses that Lexis influences his own ideas and believes that, in spite of some essential reservations about the notion of probability, Lexis is more fruitful with respect to the notion of probability and better suited to the principles of proper induction than the Pearsonian direction.

Bortkiewicz thought that Keynes judged quite favourably his papers on the theory of dispersion which had appeared a quarter of a century ago and even earlier. As to his law of small numbers, he naturally thought that it had hardly proved anything except that the Lexian criterion of stability is not applicable to the case of rare events.

Blaschke [1898] had also pronounced that opinion. It would have been correct had the small number of the occurrences of an event led to the impossibility of the coefficient of dispersion to exceed unity considerably, because of a purely arithmetical reason having nothing in common with probability theory grounds. But this is not so. Definite conditions which can only be formulated in the language of the theory of probability should have been added for that coefficient to remain in the vicinity of unity. One and only one example can explain it.

It has to do with the blowing up of steam boilers. Prussian statistics [*Jahrbuch* 1910, p. 136] shows under two different headings the yearly number of such accidents and the number of the workers who died because of them. For 1890-1909 the mean yearly numbers were 3.3 boilers and 1.8 killed. The second number is smaller but the coefficient of dispersion for the blowing up of the boilers is 0.86, whereas for the killed, on the contrary, 1.67. In the second case the mean error of the coefficient is 0.16, so that it is out of the question to say that the difference equal to 0.67 [ $1.67 - 0.86 - 0.26 = 0.65$ ] was random. It is rather explained by *acute solidarity of separate cases*.<sup>17</sup> *Sapienti sat* [Lat. 'sufficient for the clever'].

In spite of his unsuccessful, in Bortkiewicz's opinion, criticism of the law of small numbers, Keynes did not deny completely his interest in it.

<sup>17</sup> Why only one mean error is sufficient? What about the error of the first number? Mean error apparently meant *mean square* error. Lexis [1876, p. 214] defined the latter term but called it *mean error*. Elsewhere Bortkiewicz [1931, p. 9] applied the three sigma test. Bortkiewicz introduced acute solidarity earlier (1898) and Chuprov (1905) applied that term later.

He (pp. 403, 404) estimated Bortkiewicz's later work in mathematical statistics essentially otherwise and argued that *Bortkiewicz does not get any less obscure as he goes on*. "Instead of clearing up a very simple matter, I [Bortkiewicz] have befogged it with a profusion of mathematical formulas and new technical terms".<sup>18</sup> *Like many other students of Probability he is eccentric, preferring algebra to the earth*.

8. Keynes justifies this rebuke in a few marginal remarks about Bortkiewicz's paper [1918], "I have chosen its theme just as my methods, to say the truth, did not delete comma out of the blue." There are many indications in the special literature about whether homogeneity influences the stability of frequencies and how the greater or lesser homogeneity of a statistical group is here manifested.

Bortkiewicz decided to ascertain this problem by statistical data. As criteria of the degree of stability he had at his disposal the coefficient of dispersion and, since it is independent from the width of the field of observation, the essential component of the fluctuations. However, with respect to the second criterion it was necessary to show how to calculate it, since Lexis sometimes applied an inconvenient method of determining it.

Then Bortkiewicz established a criterion for the degree of the homogeneity of a statistical group and explained that, if a magnitude cannot be quantitatively estimated, we may still be sure that, in accordance with that criterion, a total group can never show a higher degree of homogeneity than the mean of its partial groups. Indeed, the population of Germany, for example, is in every possible sense less homogeneous than the populations of its separate regions on average.

Bortkiewicz have thus prepared the ground for a study of the relation between homogeneity and stability and successfully showed by a series of examples that the coefficient of dispersion for a total group, although higher than for the mean of separate groups, was not however as high as was possible to expect because of the larger number of its separate cases as compared with the partial groups.

In other words, it happened that the essential component of the fluctuation for a total group was lower than for the mean of the partial groups. According to the above, this means no more and no less than a combination of lesser homogeneity with a higher stability and vice versa. This was required to be explained and the author managed to establish a mathematical connection between the component of the fluctuation of a total group and the components of the partial groups.

<sup>18</sup> The author copied the next sentence from Bortkiewicz [1931/108, pp. 19-20] where it was naturally written in English. Bortkiewicz [1923/93] provided both neighbouring sentences in their original English.

9. The formula which expresses that connection includes a factor composed after the coefficient of correlation and called [by Bortkiewicz] the coefficient of syndromy.<sup>19</sup> It shows the measure of the mutual correspondence of the appropriate statistical series composed of separate partial groups. When the correspondence is absolute (isodromy) this coefficient is 1; when it is more or less considerable (homodromy), it is contained between 0 and 1; if there is no correspondence at all (paradromy), the coefficient is 0; and, finally, when the processes, which are described by the series, are proceeding antagonistically (antidromy), the coefficient of syndromy is negative.

The smaller the coefficient of syndromy, the more the essential component of the total group relative to the mean value of that component decreases for the partial groups which it equals in the case of isodromy.

Thus the cause of the established mutual relations between homogeneity and stability is that isodromy never really occurs and invariably the other forms of the syndromy take place. If this explanation is correct, then, the stability of the total group ought to be essentially higher than for the partial groups if it is composed of absolutely incompatible parts.

Indeed, in this case the most likely will be paradromy or even antidromy, and this really happened for the statistics of marriages in 1899-1908 in six cities, Barcelona, Birmingham, Boston, Leipzig, Melbourne and Rome taken together. The essential component of the fluctuation was so small compared with the data from the same cities taken separately, that the coefficient of dispersion for the artificially created total group was lower than the mean of the indicated parts. At the same time according to the Lexian pattern of probabilities changing serially, it should have been considerably higher than for that mean.

Indeed, Lexis derived the relation between the value of the coefficient of dispersion and the number of the separate cases (trials) for (the hardly occurring) isodromy.<sup>20</sup> Bortkiewicz supposed therefore that his study is a contribution to the theory of stability of statistical series and that, in spite of Keynes' remarks, it is exactly in the direction of a more accurate understanding of reality.

10. At the end of the paper Bortkiewicz indicated that the statement about the antagonistic relation between homogeneity and stability had

---

<sup>19</sup> Five new terms. Bortkiewicz introduced them previously [1918/68, pp. 42-43] and once more afterwards [1931/108, p. 10] but this author doubts that anyone else picked them up.

<sup>20</sup> The author has not found that derivation.

been for a long time actually applied in the practice of insurance. It became known that to ensure the calm existence of an insurance establishment, the stability of the numbers in which its activity is represented is naturally needed which was assisted by the biggest possible breakdown of insurance according to territorial or other conditions, rather than its concentration in a small district and a small variety of the ensured risks.

Keynes however thinks that this argument is only an example provided earlier by Bortkiewicz of the difference between the general probability  $p$  and its components, separate probabilities  $p_1, p_2, \dots$ . This is what he said (p. 403):

*If we are basing our calculations on  $p$  and do not know  $p_1, p_2, \dots$ , do not italicize 1 and 2 then these calculations are more likely to be borne out by the result if the instances are selected by a method which spreads them over all the groups 1, 2, ... than if they are selected by a method which concentrates them on group 1. In other words the actuary does not like an undue proportion of his cases to be drawn from a group which may be subject to a common relevant influence for which he has not allowed.*

*If the à priori calculations are based on the average over a field which is not homogeneous in all its parts, greater stability of result will be obtained if the instances are drawn from all parts of the non-homogeneous total field, than if they are drawn now from one homogeneous subfield and now from another. This is not at all paradoxical. Yet I believe, though with hesitation, that this is all that von Bortkiewicz elaborately supported mathematical conclusion amounts to.*

Let us suppose that, for example, in the case of fire insurance we deal with two types of buildings, dwellings and factories with differing risks of fire which was not however taken into account. The premiums would have rather depended on the existence of a definite ratio of those risks to each other. Then, according to Keynes, insurance of both types of buildings covering the entire considered period would have assisted the greater stability of the results of insurance, rather than the insurance of only one of the types of the buildings and only for a year, then only of the other type for the next year etc.

This is really so (and not at all unusual)<sup>21</sup>, but it is irrelevant to Bortkiewicz's statement about the antagonistic behaviour of homogeneity and stability. For a connection to this thesis we should rather contrast the following two cases. In the first one, both types of the buildings are insured for some years with a definite relation [of risks] to each other

<sup>21</sup> This statement is not clear enough.

(which only undergoes normal random oscillations). If the risks for both types of buildings are not too considerably different, a scale of premiums [introduction of differing premiums] for them will but little change anything. Such action increases the stability of the insurance activity.

In the other case only one type of the buildings is insured for a number of years and the premiums correspond to the risks of fire. Bortkiewicz's statement would have meant that the first case with a lower homogeneity ensures a higher stability compared with the possibility that for both types of buildings, risks change from year to year but in such a way (normal random oscillations are naturally not considered here) that these changes at least partially compensate each other.

Indeed, nothing except such a compensation takes place under homodromy, paradromy or antidromy. In addition, the higher stability in the first case was not connected with the condition that the mean risks for the two types of buildings over all the period are different. Even if not different, those types do not quite correspond to each other in respect of the yearly oscillations of the risks (apart from those normal random), a higher stability will occur in the first case.

The same conclusion followed from my example of 1918: the difference of the mean probabilities in separate partial groups was the factor which should be taken into account although the essence of the problem did not change.<sup>22</sup>

11. How could it happen that Keynes did not understand not just some part of my paper, but all of it, and to such an extent? As far as the author [Bortkiewicz] understands the only explanation which has a higher (although unmeasurable) probability is that the materials from which he issued were too extensive. Already the number of the sources which he had looked through excludes the possibility of a balanced scrupulous study of each of them. Moreover, the matter concerns a contribution which studies somewhat complicated connections and is not as easily coped with as with most materials, although not absolutely inscrutable.

Such particulars are hardly significant compared with the essential virtues which isolate his contribution exactly as critical information about the merits of other authors. Sometimes Keynes refers to German authors just for fun, as for example to Bobek [1891]. In his barely known textbook he [Keynes 1921, p. 383] calculated the probability of invariable sunrises during the next 4000 years and concluded that it was only  $2/3$ <sup>23</sup>.

However Keynes provides plenty of references to German authors and justly recognises the success of the German spirit and zeal in the

<sup>22</sup> Later (1931/108) Bortkiewicz better described the example concerning insurance.

<sup>23</sup> Keynes had provided no details but anyway Bortkiewicz was wrong: the sunrise problem became classical and there was nothing funny about referring to Bobek.



philosophical, mathematical and applied theory of probability. He found long-forgotten books, for example the contribution of Kahle [1735], a teacher of public and church laws, and praised them.

Keynes' erudition in the multi-branched field of probability is surprising. In this respect he rivals Chuprov. Incidentally, he characterizes Chuprov as an intermediate link between the *German* and the *English* school and praised him once more although without being acquainted with his main work (1909)<sup>24</sup> since it was published in Russian.

When recalling that Keynes according to his main position is a professor of national economy, and a publicist and politician known to the whole world, we cannot refuse him our eminent recognition of such a comprehensiveness of the highest grade.<sup>25</sup>

## References

Abbreviation: JNÖS = *Jahrbücher f. Nationalökonomie u. Statistik*.

Bernoulli J., 1685, *Parallelismus logici et algebraicim*, *Werke*, Bd. 1, Basel, 1969, pp. 265-272.

Blaschke E., 1898, *Das Gesetz der kleinen Zahlen*, Monatshefte f. Math. u. Phys., 9. Jg., pp. 39-41.

Bobek K.J., 1891, *Lehrbuch der Wahrscheinlichkeitsrechnung*, Stuttgart.

Boole G., 1862, *On the theory of probabilities. Studies in Logic and Probability*, London, 1952, pp. 386-424.

Chuprov A.A., 1905, *Die Aufgaben der Theorie der Statistik*, Schmollers Jahrb. f. Gesetzgebung, Verwaltung u. Volkswirtschaft im Dtsch. Reiche, Bd. 29, pp. 421-480 (quoted from its Russian translation of 1960).

Chuprov A.A., 1909, *Ocherki po Teorii Statistiki (Essays on the Theory of Statistics)*, Moscow, 1959.

Cournot O., 1843, 1984, *Exposition de la théorie des chances et des probabilités*, Paris, S, G, p. 54.

Couturat L., 1901, *La logique de Leibniz*, Paris.

<sup>24</sup> That contribution (1909) was written by the yet non-mathematically minded Chuprov (a mathematician by education) and my opinion about it is quite negative, see at least Sheynin [2011, pp. 124, 142 and Note 14.10 on p. 172]. Later in life Chuprov hardly referred to it and Chetverikov, Chuprov's closest student, testified that sometime after 1910 Chuprov refused to reprint his contribution; its edition of 1959 appeared long after his death.

Chuprov as an intermediate link: Keynes apparently thought about Chuprov's efforts to unite these schools.

<sup>25</sup> Here is one more criticism pronounced by that universally recognized person [Keynes 1921/1973, p. 440, Note 2]: *The mathematical argument is right enough and often brilliant. But what it is all really about, and what it really amounts to and what the premise is, becomes increasingly perplexing to decide.*

- Czuber E., 1910, *Wahrscheinlichkeitsrechnung und ihre Anwendungen*, Bd. 2, Leipzig (New York 1968).
- Edgeworth F.Y., 1911, *Probability*, Enc. Brit., vol. 22, pp. 376-403; *Writings in Probability, Statistics and Economics*, vol. 1, Cheltenham, 1996, pp. 125-152.
- Ellis R.L., 1842, *On the foundations of the theory of probabilities*, Trans. Cambr. Phil. Soc., vol. 8, 1849, pp. 1-6; *Mathematical and Other Writings*. Cambridge, 1863, pp. 1-11.
- Fries J.F., 1842, *Versuch einer Kritik der Principien der Wahrscheinlichkeitsrechnung*. Braunschweig
- Gauss C.F., 1863-1930, *Werke*, Bde 1-12, Hildesheim, 1973-1981.
- Gini C., 1946, *Gedanken zum Theorem von Bernoulli*, Schweiz Z. f. Volkswirtschaft u. Statistik, 82. Jg., pp. 401-413.
- Helm G., 1902, *Die Wahrscheinlichkeitslehre als Theorie der Kollektivbegriffe*, Annalen der Naturphilosophie, Bd. 1, pp. 374-381.
- Jahrbuch*, 1910, *Statistisches Jahrbuch f. den Preussischen Staat*.
- Kahle L.M., 1735, *Elementa logicae probabilitium*, Halle.
- Keynes J.M., 1921, *A Treatise on Probability. Collected Works*, vol. 8. London, 1973.
- Kries J. von, 1886, *Die Principien der Wahrscheinlichkeitsrechnung*. Freiburg i/B. Tübingen, 1927.
- Kries J., von, 1888, Über den Begriff der objektiven Möglichkeit und einige Anwendungen desselben, Vierteljahrsschrift wissenschaftl. Philos., Bd. 12, pp. 179-240, 287-323, 393-428.
- Kries J., von, 1916, *Logik*, Tübingen.
- Laplace P.S., 1812, *Théorie analytique des probabilités. Œuvres complètes*, vol. 7. Paris, 1886.
- Laplace P.S., 1814, *Philosophical Essay on Probabilitie*, New York, 1995 (transl. A.I. Dale).
- Lexis W., 1876, *Das Geschlechtsverhältnis der Geborenen und die Wahrscheinlichkeitsrechnung*, JNÖS, Bd. 27, pp. 209-245, reprinted: Lexis (1903, pp. 130-169).
- Lexis W., 1879, *Über der Theorie der Stabilität statistischer Reihen*, JNÖS, Bd. 32, pp. 60-98, reprinted: Lexis (1903), pp. 170-212.
- Lexis W., 1903, *Abh. zur Theorie der Bevölkerungs- und Moralstatistik*, Jena.
- Mill J.C., 1843, *System of Logic*, London 1886; many more editions, e.g. *Collected Works*, vol. 8, Toronto 1974.
- Mises R. von, 1928, *Wahrscheinlichkeit, Statistik und Wahrheit*, Wien 1972.
- Sheynin O., 2011, *Chuprov: Life, Work, Correspondence*, Göttingen.
- Sheynin O., 2017, *Theory of Probability. Historical Essay*, Berlin, S, G, 10.
- Stumpf C., 1899, *Bemerkung zur Wahrscheinlichkeitslehre*, JNÖS, Bd. 17 (72), pp. 671-672; Bd. 18 (73), p. 243.
- Stumpf C., 1893a, *Über den Begriff der mathematischen Wahrscheinlichkeit*, Sitz.-Ber. Bayer. Akad. Wiss. München, Philos.-Philolog. und Hist. Kl., Jg. 1892, pp. 37-120.
- Stumpf C., 1893b, *Über die Anwendung des mathematischen Wahrscheinlichkeitsbegriffe*, pp. 681-691.
- Todhunter I., 1865, *History of the Mathematical Theory of Probability*, New York 1949, 1965.
- Venn J., 1866, *Logic of Chance*, London 1888; 1962.
- Wagner K., 1898, *Das Problem vom Risiko in der Lebensversicherung*, Jena.