BULLETIN

DE

L'INSTITUT INTERNATIONAL

DE STATISTIQUE

TOME XXII - 1ère LIVRAISON

ROMA
PROVVEDITORATO GENERALE DELLO STATO
LIBRERIA
1926

TROISIÈME PARTIE.

ANNEXE B.

THE REPRESENTATIVE METHOD IN PRACTICE

By Adolph JENSEN

I. — Introduction.

« The representative method »–it sounds so ceremonious, and yet it is true that this method is one of the most common means of comprehension, not only in scientific research but in life generally.

All our knowledge is piece work. We form practically the whole of our image of the world by means of a series of partial observations, and from a generalisation of these we arrive at results which, at any rate to a certain extent, we regard as being of general validity.

Whether these results are correct or incorrect depends upon a number of circumstances, but not least upon whether we have succeeded in rightly valuing the representative character of the observations. In the face of the manifold phenomena which go to make up our daily life experience, in this respect as in so many others, is an excellent teacher. Without being conscious of it we are schooled from the first years of childhood in making daily use of the representative method with greater and greater certainty.

From the daily life the application of the method has been unobservedly transferred to investigations in support of practical work and finally to most domains of scientific research.

The exploration of the earth by means of geological, botanical and zoological investigations for the most part rests upon samples. We have acquired our knowledge of the oceans by means of deep-water investigations which we presume are representative. Our ideas of meteorological conditions depend upon a systematic treatment of single observations. And so on. To name only one more prominent example: all modern biological science is, so to say, built up upon representative investigations.

In the sphere of statistics the representative method has also, in a more or less perfect manner, been used for centuries. This was the case when men like Will. Petty, Gregory King, or Charles Davenant groped their way through the sparse, casual and uncertain

numerical material which was at the disposal of statisticians about two hundred and fifty years ago ; or when French votaries of « Political Arithmetic » as, for instance, Messance and Moheau, a hundred years later calculated the probable distribution of the population according to sex, age and conjugal position ; or when Lavoisier calculated the production and consumption of foodstuffs in France. After all, what was «Political Arithmetic » other than a groping and awkward use of methods which, evolved in more systematic forms, were later on adorned with names such as « Sampling » and « Representative Investigations ? ». When Vauban, for instance, examined one or two small territories in various provinces and from these « samples » drew far-reaching conclusions as to the country as a whole, he in reality made use—although in an imperfect way—of the representative method. The loose examples indicated here could be augmented by many others (1). To describe merely the main lines in the development of the representative method would, in reality, be the same as to try to confine an essential portion of the history of statistics in a nut-shell. We will refrain from doing so, and only further call to mind the numerous attempts to calculate the population at a time when the actual census of population was unknown. The most classical example of this is doubtless the calculation of the population of France which Laplace undertook in 1802 on the basis of the annual number of births in the immediately preceeding period. There is so much the more reason to call attention to this investigation, as Laplace went a considerable step further than most of his predecessors by determining the limits of error in the result (2).

Despite the fact that the primitive calculations of « Political Arithmetic » in many cases were proved to have given comparatively good results, they were for a time often viewed with a pitying shrug of the shoulders. It was then that statistics had advanced so far that, with regard to certain fundamental circumstances, they could build upon complete counts. Population censuses in particular put the statisticians on a firm footing, and they gained a feeling of certainty which they had not formerly known. Thus it is not to be wondered at that there arose a sort of dogma that the most extreme degree of accuracy must be demanded of statistical results. With the population censuses and other great collections of numerical material the bureaux of statistics were developed, and statistical work became a

---

(1) See, for instance, H. WESTERGAARD, *Statistikkens Teori i Grundrids*, Köbenhavn, 1915, pp. 27-59.

(2) LAPLACE. *Théorie analytique des probabilités*, 2ième édition, 1814, p. 391.

kind of profession which on the one hand favoured the tendency of the demand for accuracy to develop in the direction of pedantry, but on the other hand without doubt made a rational development of statistical methods possible.

In the latter half of the 19th Century especially, official statistics had been extended to comprise a number of phases in the life of the community as to which one was formerly quite in the dark. But many of the new problems were of such a nature that it was impossible to procure a complete material, and thus one was brought face to face with the task of having to find out the probable composition of a mass without being able to bring all the units of the mass under observation. Naturally, attempts have been made to solve this problem in a variety of ways.

In the so-called *Inquiry Method* (« enquêtes ») the direct observation of the units of the mass were replaced by competent estimates of the general composition of the mass, most often for the purpose of ascertaining the typical features of the mass, with particular regard to those characteristics about which knowledge is desired.

Another principle of investigation is brought into application in the *Monographic Method*. This, too, aims at elucidating the typical features, but whereas the Inquiry Method must, according to its nature, to a certain degree build upon a certain subjective estimate, the Monograph builds upon objective observations of a strictly limited number of units, which are subjected to a careful examination and often made the objects of very detailed description.

But neither the Inquiry Method nor the Monographic Method are statistics in their real sense. The Monograph may even be said to be the diametrical opposite to statistics, of which the distinctive feature is that it works with mass-observations. Both the Inquiry and the Monograph, and the methods related to them, can in many cases suffice very well; indeed, in certain respects they may also give a more complete and better outlined picture than can be obtained by means of a purely statistical investigation. But knowledge of the composition of the mass in detail can only be attained by means of statistics, that is to say, by means of objective observations under such circumstances that the units of the mass, of every kind, come under observation regardless of whether this or that kind coincides with the typical or the generally appearing, or whether it perhaps only appears rarely or as an exception.

Now as the same information as can be obtained by means of a complete statistical investigation of a mass material can, in many cases, be obtained with sufficient accuracy by means of an investi-

gation of a part of the material, it is not to be wondered at that partial investigations in our time have been extensively used in the most varied domains : in administrative statistics and transport statistics (postal and telegraph despatches), in agricultural statistics (crop forecasts and crop returns, live stock statistics), in price statistics (index numbers), in labour statistics (wages and working hours, etc.), in statistics on conditions of social life (consumption, food, housing conditions, etc.), and to some extent in demographic statistics (especially the border domains between these and social statistics).

The representative method can, however, be used in many different ways; it is precisely the object of this Memorandum to show, by means of a number of examples, how the method can be adapted to the various problems which present themselves in practice. But before we proceed further, we must first submit a few observations as to the proper limitation of the term « Representative Investigation ».

In statistical terminology the expression « Representative Investigation » is sometimes given a wider and, at other times, a narrower significance. Sometimes partial investigation has been described as « representative », simply because it takes the place of (and thus represents) a complete investigation — especially when the portion examined is comparatively comprehensive. As an example of this may be mentioned the comprehensive enquiries made by the British Board of Trade in 1886 and 1906-07 into the earnings and hours of labour of workpeople, and the investigation made by the French Statistique Générale into the size and composition of families belonging to the employee and working-classes (« Statistique des familles en 1906 »).

In other cases, a partial investigation has been described as representative simply because it is *presumed* that its results give a correct reflection of the whole. As an example may be mentioned the investigations which the British Ministry of Labour make every month into the state of employment in the various industries, the results of which are published in the Ministry of Labour Gazette. The material for these employment returns is not selected according to any principle ; information is obtained from as many establishments as possible. Nevertheless, these returns are described as representative, being put forward as affording at least some indication of the state of employment in the various industries as wholes.

An example of quite another kind is the following : In Switzerland the « Office fédéral du Travail » has for a number of years drawn up statistics of wages for workers who have been the victims

of accidents. These statistics have been called representative, naturally from the point of view that « chance » determines whether a worker is injured during the performance of his duties, and that the injured workers may therefore be assumed to form a sample which is representative in proportion to the total number of workers. A closer consideration will, however, show that one cannot be certain that this presupposition will always hold good. It is, for instance, reasonable to suppose that accidents mostly occur to the young (inexperienced) workers and perhaps to the oldest, and even if both the youngest and the oldest have perhaps lower wages than the average, it is by no means certain that the factors named nullify each other's effects. In the reports on the results of these investigations a number of other factors, besides those herein referred to, are pointed out, and these make it doubtful whether the investigations can be regarded as representative or not (1).

Even if one can, with some justification, call investigations such as we have referred to above « representative », we prefer, however, to reserve this expression for investigations which are really made by the application of the representative method, that is to say, where in some way or other care has been taken that that part of the whole which is the subject of direct observation is representative in proportion to the whole. In the following we will confine ourselves solely to such investigations as are « representative » in this sense.

The following pages contain a number of examples of representative investigations in various spheres and from various countries. The choice of examples has not been directed by any principle; both good and bad have been included, as the object is to give an illustration of the manner in which the representative method has actually been applied in practical statistics.

The examples taken are arranged according to the procedure which has been used in order to produce a representative sample.

To the purely objective descriptions of the methods used (Sections II-VI), various observations of a general and critical nature have been added (in Sections VII-VIII).

As an appendix to this survey of the application of the representative method in practice, a list is added of books and treatises, etc. in which this and similar methods are treated from a theoretical point of view (Section IX).

---

(1) See particularly : *Informations de Statistique Sociale*, Year 1, No. 3, pp. 6-8 (Berne, 1923).

II. — Random Selection of Units.

In this section a number of examples of representative investigations will be given, where the representative character of the sample is ensured by selecting from the whole mass of units a certain fraction at random. In these investigations it was thus a presupposition that each single unit in the mass has an equal chance of inclusion in the sample.

The representative method may often be used with advantage where, after a summary treatment of a mass-material has been undertaken it is desired to extend enquiry with regard to certain points. As an example of this the following is quoted.

In connection with the Danish census of population in 1901 an investigation was made regarding the existing marriages. The information in question was given in special columns on the census schedules. It was the object of the investigation to show the difference between various groups of the population with regard to marriage frequency, marrying age, duration of marriage, number of children born in marriage and infant mortality in marriage. Having regard to the funds which were available, the investigation had, however, to be limited to 21 occupation groups. The material covered as a rule for each of these occupations from 1000 to 4000 marriages; in three occupations only, viz. farmers, small holders and day-labourers on farms, were the figures very high, 50,000, 47,000 and 28,000 respectively. For these three occupations one simply selected one fifth of the marriages, the sample being formed by mechanically extracting every fifth marriage of the three groups named (1).

As another example of a similar kind the following is advanced:

When the Norwegian census was taken in 1900 it was desired to make a separate investigation into the size and composition of households combined with the occupation of the head of the family, and also with regard to the position of the various persons in the household combined with their sex and age, etc. In this very far-reaching special investigation, which was made concurrently with the more summary investigation of the whole material of the census, the representative method was applied in the following manner:

The census material was in the form of lists, one for each family.

---

(1) *Ægteskabs-Statistik.* (Statistiske Meddelelser, Series 4, Vol. 18, København, 1905).

As it was recognised as being impossible to include all the lists in the investigation, there was selected, in the towns, every tenth, and in the rural districts every twentieth, list. In the first census district were taken, for example, in the towns the lists numbered 1, 11, 21, etc., in the rural districts the lists numbered 1, 21, 41, etc.; in the second census district were taken, in the towns, lists numbered 2, 12, 22, etc., and in the rural districts lists numbered 2, 22, 42, and so on. In this investigation, where the unit was not a person, but a household, the sample was thus selected at random by quite a mechanical counting out. After the investigation it turned out that the sample, with regard to various ordinary demographic characteristics very closely agreed with the whole (1).

The examples given in the foregoing refer to investigations of a purely demographic kind. The random selection of units has, however, also been applied to enquiries on conditions in social life. In these domains it will certainly often be difficult to fulfil all the requirements of a fully reliable random selection. As an example it may be mentioned that in most investigations regarding housing conditions and rent it has been necessary to select a sample comprising certain streets or certain quarters of a town. As examples of the application of random selection in enquiries into housing conditions of less well-to-do families there may, however, be mentioned two investigations of this kind, made in Gothenburg in 1911 and 1921. The procedure was as follows:

To help in the selection of the sample lists were prepared of all the real properties in the town, showing the number of rooms in each single dwelling. Based upon these lists a slip was made out for each of the properties in which there were small dwellings (with less than three rooms) showing the address of the property and the number of small dwellings in it. The slips were mixed together in an urn and then drawn one by one, until for each of the districts in the town about a fifth of all the small dwellings had been drawn. The properties so drawn were investigated by visits to them, and the investigators were instructed not to pass over any of the small dwellings in the properties without most urgent cause (2).

A similar method to that just described was used in an investigation into housing conditions in small dwellings in Christiania in 1913-14. In the Statistical Office there were lists of all the pro-

---

(1) *Familiehusholdningernes Sammensätning.* (La composition des ménages) Norges officielle Statistik. Series 4. No. 82, p. 90.

(2) Information kindly furnished by Dr. J. GUINCHARD.

perties showing the size of all dwellings in them including the number of rooms. In each street in the whole town every third house was selected if in the house there were dwellings of not more than two rooms. It turned out that this method led to a comparatively very equal representation both of the various parishes and of the various dwelling groups. The representative character of the sample was further tested in various ways; it proved, for instance, that both the rent and the average number of persons per dwelling of the various types were very nearly equal in those specially investigated and in other dwellings. The investigation itself was made by actual visits to the houses, the investigators being furnished with complete instructions. Of the dwellings selected for investigation, 0.8 per cent had to be rejected, one of the reasons being that the occupiers, despite repeated visits, could not be found at home (1).

It is clear that there will be particular occasion for applying the representative method in cases where the information required can only be obtained by personal interviews with a large number of people. Such oral enquiries take up a lot of time and must necessarily be made by a comparatively small number of trained and properly instructed enquirers. Many investigations of this kind would therefore be impracticable if one could not be content with investigating a fraction of the units. In the following the procedure in an interesting enquiry of this kind will be commented on.

For the purpose of getting a general picture of the classes of persons who make up the numbers of claimants to unemployment benefit in Great Britain, the Ministry of Labour made in November, 1923, an investigation into the personal circumstances and industrial history of 10 000 individuals selected from among the 1 054 000 persons who at that time were being returned as « insured workpeople unemployed » at the Employment Exchanges of Great Britain, not including Branch Offices.

The enquiry was spread over the whole of the 413 Employment Exchanges. The procedure followed as regards the allocation of the work of the enquiry was to require from each Exchange manager a detailed report concerning a specified number of individuals on his register. This number was fixed, on the average, roughly in the proportion of 1 to 100 of the total number of claimants at that Exchange; but departures from this ratio were necessary in many cases in order that an undue amount of work might not be thrown on the

---

(1) *Beboelsesforholdene i smaa lejligheter i Kristiania*, 1913-1914. (Condition d'habitation dans les petites logements) Christiania, 1915.

managers of the largest Exchanges, and that the reports might, nevertheless, be distributed over the whole country.

The procedure adopted for the selection of the individuals to be interviewed was to require managers to « tab » every 100'th claim (the 100 being subject to the above-mentioned departures) in their « Live Claims » files. Claims are kept in the files in occupation order, and in order that there should not be an undue preponderance of any one occupation among the persons selected, the selection of the first claim to be « tabbed » was laid down for each Exchange, the starting-point being thus made to vary from Exchange to Exchange, and the starting-points were distributed as evenly as possible over the whole occupational classification. This procedure was followed separately for men, women, boys and girls.

Each of these « tabbed » claims was regarded as making the centre of a small group of claims each of which should furnish one claimant. Arrangements were made that the first person in any group coming to the Exchange should be interviewed. This group arrangement was adopted in order to ensure that the work of interviewing should not be held up by the non-appearance of a selected candidate. For each of the interviewed claimants a schedule was filled in by the Exchange manager, and all the schedules were forwarded to the Ministry of Labour for analysis (1).

To some extent, it was possible to check the results of this 1 per cent sample by comparing some of the results with the results of two other investigations made on larger samples at other dates in the same year. On a certain day in January, 1923, every third claim in the live claims files at Employment Exchanges had been examined and all relevant information abstracted for purposes of statistical analysis (2). Further, in order to obtain precise information of a very limited kind concerning the number of claims made over a long period (as distinct from the number of claims current on a particular day) an investigation was made towards the end of 1923. An examination was carried out of 10 per cent of all the claims made during a special period, in the course of which information as to the ages of claimants whose claims were « live » at the end of the period was obtained and tabulated (3).

---

(1) *Report on an Investigation into the Personal Circumstances and Industrial History of* 10,000 *Claimants to Unemployment Benefit.* London, 1924. (Summarised in the « Ministry of Labour Gazette », March, 1924).

(2) *Analysis of Claims to Unemployment Benefit.* « Ministry of Labour Gazette », November, 1923.

(3) *Unemployment Insurance : Analysis of Claims to Benefit during the Fourth Special Period.* « Ministry of Labour Gazette », July, 1924.

It is evident that the results obtained by the two above mentioned large-sample enquiries were not strictly comparable with the results of the 1 per cent sample enquiry. None the less, a comparison of certain particulars (age and marital distribution and dependents' benefit) in the three samples demonstrates that the 1 per cent sample was nowhere very wide of the mark. Except where the absolute numbers involved were very small, the correspondence between age and marital distribution as shown by the three samples is remarkably good. Mr. John Hilton, Director of Statistics at the Ministry of Labour, has expressed the opinion that the 1 per cent sample has answered most of its purposes quite as well as a 10 per cent or 33 per cent enquiry would have done (1).

Where the random method is used, it is usually in such a way that a certain fraction of the total mass is selected according to some mechanical principle or other; by multiplying the figures in the result by the denominators of the fraction, one arrives at the figures which apply to the whole mass. We have, however, already referred in the foregoing to an investigation where the representative method was only applied to certain groups in the whole, whereas the other groups were subjected to a complete investigation viz., the investigation into the difference between various groups of the Danish population with regard to marriage frequency, number of children born in marriage and infantile mortality. The basic idea of such a division of the investigation can now be carried forward in that way that the relative size of the sample is fixed differently for the various parts of the whole. If, for instance, the mass consists of three groups, A, B and C, of which A is completely investigated, whilst B is investigated by means of a sample of one-fifth, and C by a sample of one-tenth, the total result will be arrived at by the formula : $a + 5b + 10c$. The following is an example of such an investigation.

On the basis of the municipal rating assessments for 1913-1914 an investigation was made in Norway into the distribution of the population according to income and fortune, combined with occupation. As far as most of the assessment districts were concerned, all rate-payers were included in the investigation, but for the large towns and some of the largest circuits in the country the investigation was only complete with regard to the largest rate-payers, whilst the other rate-payers were investigated representatively, the fraction according to which the sample was selected in some cases being one-third, in others one-fifth, and in others one-tenth (2).

(1) John Hilton, *Enquiry by Sample, an Experiment and its Results.* (Journal of the Royal Statistical Society, July, 1924).

(2) *Norges officielle Statistik*, VI, 57. Christiania, 1915.

In the years 1912-1914 Professor A. L. Bowley directed some enquiries into the economic conditions of the working-class in five English towns. The first of these investigations was made in the borough of Reading, in the autumn of 1912. A sample was selected from the whole of the borough as follows: one building in ten was marked throughout the local directory in alphabetical order of streets, making about 1950 in all. Of these about 300 were marked as shops, factories, etc., institutions and non-residental buildings, and about 300 were found to be indexed among principal residents, and were so marked. The remaining 1350 were working-class houses, and a number of volunteers set out to visit every one of these. It was presently found that the scale taken was beyond their powers, and it was decided to take only one house in 20, rejecting the incomplete information as to the intermediate tenths. The visitors were instructed never to substitute another house for that marked, however difficult it proved to get information, or whatever the type of house. In the end the enquiry-committee failed to learn anything as to 32 households out of 677, and substituted for these 32 of the surplus tenths, without, so far as could be judged, introducing any bias. On examination, it proved that in 55 of the 677 the occupier was above the working-class (clerk, traveller, shop manager, etc.), and these were excluded from the investigation (1).

This investigation was followed in 1913 by similar investigations into the economic conditions of the boroughs of Northampton and Warrington, and the urban districts of Stanley. Approximately one working-class house in 23 was visited at Northampton, one house in 19 at Warrington, and one house in 17 at Stanley. For each town a list of all houses, as given in a directory for Northampton, and in the assessment lists and burgess rolls elsewhere, was obtained and the sample was taken without reference to anything except the accidental order (alphabetical by streets or otherwise) in the list. A similar investigation was conducted in Bolton in 1914. It was decided to take a sample of 1 in 10 instead of 1 in 20 in the other towns, and, since Bolton is two or three times as big as Northampton, Warrington or Reading, and its non-working population is relatively small, the sample included 3650 households, as compared with 600 to 700 in the other towns.

In all these working-class investigations a series of interesting particulars was procured. On a prepared card the investigators inserted

---

(1) A. L. Bowley. *Working-Class Households in Reading.* (Journal of the Royal Statistical Society, June, 1913).

details of the age, sex, and membership of the family, the occupation of those employed, the apparent income, and the housing accommodation. The results of the investigation showed hitherto unknown details of the variation of the constitution of the family, of the varying numbers of earners and dependents, and especially of the burden of children on various grades of income. Further it showed in what proportion of households income did not suffice for a certain minimum standard of living, and the circumstances which led to such deficiency. The results were controlled by general local and other official statistics, where these were available. The results, together with detailed information on the method and arrangement is given in a volume published by « The Ratan Tata Foundation » (University of London) (1). In this volume a special chapter (by Professor Bowley) is devoted to the criticism of the accuracy of the results.

A similar investigation was made in the same towns in 1924, the results of which are to be published this year.

To supplement the information obtained in the working-class investigations which are mentioned above, permission was obtained to have transcriptions made of certain details from the householders, schedules of the census of England and Wales of 1911. One schedule in 50 was chosen *seriatim* for Bristol, Newcastle-on-Tyne, Leeds, Bradford (Yorks.), and three London Boroughs. The results (published in «Economica», No. 2, at the London School of Economics) showed again in great detail the varying constitution of the working class family.

These results have been extensively used by persons who advocate « Family Endowment » (2).

That the representative method with the application of random selection can give good results, even when the subject of the investigation is of a fairly difficult character, will appear from the following example :

In the town of Mannheim an investigation was made in 1915 into certain economic effects of the state of war. It was desired, among other things, to know to what extent payment of house rent due had not been made ; to what extent landowners were in arrears with the payment of interest or mortgages ; finally, to what extent mortgage-debts on real property had been foreclosed and replaced by mortgages at a higher rate of interest.

---

(1) A. L. Bowley and A. R. Burnett-Hurst : *Livelihood and Poverty.* London, 1915.

(2) Information furnished by Professor Bowley.

The sample used in this investigation was thus obtained. In the directory every tenth property was selected by purely mechanical counting out and to each of these properties a question form was sent, these in 436 cases being filled up so completely that the forms could be included in the representative investigation. But apart from these 436 properties which were of use to the investigation, there were 105 properties for which the desired information for various reasons could not be procured. It was therefore necessary to ascertain whether the non-inclusion of these 105 properties would detract from the representative character of the sample selected at random.

Various tests were made to find this out. One of these tests was a distribution of both the 436 and the 105 properties according to the average number of inhabitants per room. According to these average figures the properties were divided into 6 groups, as shown below and thereafter it was calculated how many of the 105 properties would come into each of the groups if their distribution according to the density of habitation had been in the same proportions as the 436 properties. The result was as follows :

Average Number of Inhabitants per Room	Distribution of the 105 Properties	
	Actual Distribution	Calculated Distribution
Up to – 0.75	11	7.6
0.76 – 1.00	19	14.7
1.01 – 1.50	27	33.1
1.51 – 2.00	29	26.7
2.01 – 2.50	11	16.5
above – 2.50	8	6.4
	105.0	105.0

Even if there is by no means complete agreement between the two columns, it is clear, that this test was a good contribution towards ascertaining the effect of the 105 properties being rejected from the material.

A direct test of the representative character of the sample used was made in the following manner : —

From a complete dwelling-house census taken in 1920 it was known how all the dwellings in the town were distributed according

to the number of rooms. A similar distribution was made of the dwellings in the 436 properties which formed the sample; see the following table: —

DWELLINGS WITH	Total Dwellings in 1910 %	Sample in 1915 %
1 room	17.1	13.3
2 rooms	40.3	40.7
3 »	22.8	26.2
4 »	9.6	9.4
5 »	4.5	5.5
6 »	2.5	2.6
7 or more rooms	3.2	2.3
	100.0	100.0

As to the results of the investigation itself, we refer to the report issued on it, which also contains much of methodological interest (1).

III. — Random Selection of Groups.

The investigations we now proceed to consider were, like the foregoing, made by the use of the random method, but with this difference, that it was not the various units in the sample that were selected at random, but groups of collateral units. As in the previous section we shall begin with a very simple example taken from demographic statistics.

A census of population was taken in Denmark on February 1st, 1921. A month later a special occasion arose for knowing as exactly as possible the distribution of age of the population in the rural districts of Maribo County. The information was contained in the census papers, but time was so short that a complete treatment of this material was not possible. The summary number of population,

(1) Dr. Sigmund Schott: *Mietausfall, Zinsrückstand und Fälligwerden von Pfandlasten in Mannheim.* (Beiträge zur Statistik der Stadt Mannheim. 4 Sondernummer). Mannheim, 1915.

divided into sexes, on February 1st, 1921, was known, and one knew very accurately the distribution of age ten years previously, viz., on February lst, 1911. Thus there was the option of making use of this latter knowledge by dividing the population as at 1921 proportionately with the distribution of age in 1911, or undertaking a partial treatment of the material for 1921.

The latter alternative was chosen. The sample was formed by mechanical selection of each third parish in the county, and in these parishes the distribution of population was investigated according to sex and age. By comparing the results of this partial investigation with the results of the investigation of the whole, which was made later, it turned out that the differences between the figures from the complete and the partial investigation in all cases except one were within the margin allowed by the law of exponentials.

The same groups obtained from the partial investigation were experimentally calculated on the basis of the knowledge gained from the census of 1911, and it then proved that the right choice was made in proceeding with the partial investigation.

Naturally, one would not ordinarily determine the distribution of age of a population by partial investigation. In this there is such a fundamentally demographic condition that nothing ought to be neglected in order to arrive at the results with the greatest possible accuracy. One may only take the liberty of a short cut to obtain a purely preliminary result in cases like that referred to, where one is at the last extremity.

A very comprehensive manipulation of a population census material by the application of the random method has been made by Dr. Sigmund Schott of Mannheim.

This concerned the working up of the census information of 1916 for the civil population in the town of Mannheim. Briefly, the procedure was as follows : —

The whole census area was divided into 1081 census districts of various sizes. These 1081 districts were given consecutive numbers and were thereafter divided into 10 principal groups, a numerical principle being strictly followed in this division. Under the first principal group were placed all those districts whose numbers ended with 1, (viz. Nos 1, 11, 21, etc. up to 1081) ; the second main group covered districts whose numbers ended with 2 (viz. 2, 12, 22 etc. up to 1072) ; and so on. Each of the ten principal groups thus covered 108 districts.

After this were formed the so-called « working groups », each containing 11 (in some few cases only 10) districts. In the formation

of these working groups, however, the arithmetical sequence was not followed. The first working group in the first principal group comprised districts No. 1, 101, 201 etc., the second No. 11, 111, 211, etc. Thus by this means the random method was again brought into use, so that every hundred'th part, i. e., every working group, formed a mechanically selected sample of the totality.

One of the purposes of the enquiry was to find out how the people were distributed into households of various sizes, and the distribution of these people according to their position in the household. The procedure may now be sketched as follows :

One may imagine that the 1081 census districts were consecutively numbered urns, in all containing as many balls as there were people in the whole population. These balls may further be imagined as being of different colours, each colour meaning a certain kind of position in the household (housewife, daughter, servant girl, etc.); again, that each ball had a figure which showed how many people there were in that particular household. The procedure was then that the balls in urns No. 1, 101, 201, etc. were collected into one big urn, and for each of these big urns, which thus represent the working groups previously mentioned, the balls were divided according to colour and figure. The contents of the urns No. 1, 101... 11, 111,... etc. were again collected and the former calculation repeated. With regard to a series of numerical proportions the principal groups were finally gathered in twos and the working groups in fives into new groups of a higher order, so that the original tenths and hundredths calculations could be supplemented by fifths and twentieths calculations.

As to the results of this interesting investigation, Dr. Schott's own report is referred to (1).

A very simple example can be given in the following investigation, which was once made as an experiment by *Marcus Rubin*.

The number of marriages, births and deaths among the rural population in Denmark in the years 1890-1894 was calculated on the basis of a sample consisting of the first circuit in each of the 17 counties, whereafter the figures were compared with the actual figures for the totality.

---

(1) Dr. Sigmund Schott: *Das Stichprobenverfahren in der Städtestatistik.* (Beiträge zur Statistik der Stadt Mannheim. No. 34, 1917).

The result will be seen from the following : —

	A. The 17 Circuits	B. The 17 Counties	A. in percentage of B.
Population 1890	204, 702	1, 366, 364	14.98
Marriages 1890-94	13, 634	93, 462	14.59
Births »	30, 327	203, 950	14.87
Of these, illegitimate	2, 702	14, 888	18.01
Deaths 1890–94	18, 934	203, 950	15.49

From these figures Rubin concluded that the sample was sufficiently representative, except as regards the illegitimate births (1).

In vital statistics the information will often present an opportunity of applying enquiry by sample as a substitute for complete investigation. By this is not meant the working out of mortality tables, although such calculations resemble the representative investigations in that a certain sequence of observations is applied to other collections of material. But apart from this, there is every occasion to endeavour to save time and labour, when circumstances make it advisable, by simply undertaking a partial treatment of the great material which offers in vital statistics.

Here, however, the difficulty presents itself that the statistical results in this domain are often required for such small geographic areas that the figures, in spite of everything, are so small and the differences so great, that a thinning down of the totality material will make the results too uncertain.

This may be illustrated by citing an attempt which has been made in Holland.

When in Holland measures of economy had to be applied to all the Departments of State, the Central Office of Statistics tried to find out, whether it would be possible to follow the representative method with regard to mortality and causes of death, without detriment to statistics.

It was decided to examine whether it would be sufficient to take, instead of all communes of less that 5000 inhabitants, or of 5-20,000 inhabitants, $^1/_5$ of these groups. In Holland it is necessary, however to have separate mortality statistics for each province, because the death rates are very different in the various provinces.

(1) See the Danish « Nationalökonomisk Tidsskrift », 1898, p. 103.

As regards the principles applied in the experiment, the following information has been given by Mr. Methorst, Director of the Central Office of Statistics.

For each province, in which the representative method is to be applied, as regards groups of communes having 5,001-20,000 inhabitants, or 5,000 and fewer inhabitants, it will be necessary to investigate whether the total of the province will not, *more than can be admitted*, deviate from the real figures, as regards :

(*a*) the mutual relations of the various age groups ;

(*b*) the mutual relations of the various causes of death ;

(*c*) the mutual relati ns of the various age groups in each cause of death.

(*d*) the mutual relations of the various causes of death in each age group.

For the determination of $^1/_5$ (20 %) as representative of the groups of communes of 5001-20 000 inhabitants, and 5000 and fewer inhabitants, it is possible to make use of different criteria, which all satisfy the demand that the choice should be left to chance only. Let us mention three of such criteria.

It is possible to let the choice depend on the letters of the alphabet (1st method), but also on the size of the successive large communes, and, in accordance with this latter method, it is possible too either to combine the larger communes with larger ones, and the smaller with smaller (2nd method), or to join larger and smaller communes belonging to the same group (3rd method), in accordance with an order to be stated beforehand.

In order to find out whether the application of the representative method could be admitted with relation to the above distinctions within the borders of the provinces, it would, strictly speaking, be necessary to have 5 experiments for each of the two groups of communes according to size, and in accordance with each of the 3 methods. Needless to say that a great many other combinations would be possible, and that these too would do just as well. Each experiment would include another $^1/_5$ part, consequently there would be 15 experiments for each group of communes. But as each of the 15 experiments of the one group of communes (5,001-20,000 inhabitants), combined with any of the 15 experiments of the other group (5,000 and fewer inhabitants) may influence in *another* way the calculation of the total for the province, it is not to be said with certainty that each « chance selection » including $^1/_5$, would be permitted, before $15 \times 15 = 225$ experiments have been made in each Province, for, even if only a few experiments should result in a greater difference

between the real, and the calculated figures for the province in total, than that which is permitted, the application of the method within the borders of the province would not be admissible for this purpose.

For each of the 225 experiments it would consequently be necessary to examine, whether the deviation of the figure calculated for the province be not too large, as regards the mutual proportions, *a* of the age groups, *b* of the causes of death, *c* of the age groups in each cause of death, *d* of the causes of death in each age group.

Mr. Methorst caused some investigation on a small scale to take place. These investigations relate to the Province of North-Brabant, in which there are 29 communes having 5,001-20,000 inhabitants, and 143 communes having 5,000 and fewer inhabitants. These two groups according to size were both of them divided into 5 sub-groups (20 %), the division being made according to the alphabetical order of the initials of the communes.

The experiment appeared, however — and this had really been expected — to be unsuccessful, the numbers being too small, and the difference between the communes too great. Mr. Methorst resumes the result of the experiment as follows : —

1. The calculated proportional figures of the age groups for the whole Province, depending on the chance selection made, are sometimes so different from reality that they rob the information of a great deal of their significance.

2. This is also the case with the calculated proportional figures of the causes of death.

3. The division according to age and causes of death in their mutual relations is made altogether useless in case of application of the representative method.

4. In case the divisions of each Province must be calculated, the application of this method does not save any time worth mentioning (1).

Using the representative method, the Austrian « Staatsamt für Heerwesen » has made an interesting investigation concerning the people belonging to the former Austro-Hungarian Monarchy who died in the war of 1914-18.

In this investigation it was not a question of ascertaining the number of dead, but the distribution of the fallen according to nationality (language) and age. As the sum total was very large, a compromise was made by examining a sample of 130,000, or about one-tenth of the whole. This sample was formed thus : of the casualty

(1) The above information has kindly been given by Mr. H. W. Methorst.

lists in the Ministry of War, eight groups were taken, distributed over various arbitrarily chosen periods from 1st August, 1914 till 19th April, 1918. The representative character of this sample was tested by the calculation of the following proportionate figures, which indicate the distribution of the fallen according to country (1):

	Whole	Sample
Austria	56.7	56.6
Hungary	40.3	40.3
Bosnia-Herzegovina	3.0	3.1
	100.0	100.0

It is rare that the representative method has been applied to investigations in the domain of pure economic statistics. We will, however, refer to one peculiar example of this:

M. de Foville repeatedly organised in the period from 1878 to 1903 representative investigations into the nationality and age of the silver coins in circulation in France. As it was impossible to collect all the coins for examination, de Foville chose the procedure of having a count made of all the coins in all public offices when business ceased on a certain day, the coins at the same time being grouped according to the country in which they were minted and according to year (2).

We will now proceed to mention some examples of the application of the representative method to the usual mass material of agricultural statistics.

In Denmark investigations regarding the utilisation of the agricultural area were formerly only made at intervals of several years, the effect of which was that the statistics of the crop yield became rather uncertain in the years for which no statistics as to the utilisation of the area were to hand.

Statistics of this kind were compiled in 1896 and 1901, but when the crop yield for 1901 was to be calculated, the area figures for the various kinds of crops were only counted for three of the 18 counties of the country. For the other 15 counties a provisional

(1) Dr. Wilhelm Winckler: *Die Totenverluste der öst-ung. Monarchie nach Nationalitäten.* Die Altersgliederung der Toten. Vienna, 1919.

(2) Bulletin de statistique et de législation comparée, publié par le Ministère des Finances, fascicules d'Octobre 1878; Août 1885; Août 1891; Décembre 1897: Mars 1904. Communication faite par M. de Foville à l'Association française pour l'avancement des sciences. Congrès de Besançon, 1893.

count of the areas in each fifth parish was therefore made and on the basis of this 20 per cent sample the figures for the totality were calculated, and these provisional figures were used as a base in the calculation of the crop yield.

The subsequent complete enumeration of the area statistics made it possible to measure the accuracy of the results of the partial investigation. The differentiations were as follows:

Wheat	2.2 %
Rye	0.2 »
Barley	0.8 »
Oats	0.0 »
Mixed Seed	4.8 »
Potatoes	1.1 »
Beet	2.5 »
Other crops	2.8 »
Hay	0.6 »
TOTAL	0.5 »

It will be seen that the differences in most cases were only small. It must also be taken into consideration that the actual changes in the utilisation of the area from 1896 to 1901 were many times greater than the differences of the partial investigation (for instance, for root crops 51 per cent, for wheat 19 per cent, etc.). The errors committed by calculating the crop yield on the basis of the figures of the partial investigation were thus insignificant compared with the errors which would have been committed by using the absolute area figures from 1896 (1).

Attempts have repeatedly been made to apply the representative method in enquiries regarding stocks of cattle.

So long ago as at the meeting of the International Institute of Statistics in Berlin 1903, the results of two such attempts were presented. One of these had been made by Dr. Kluge (2), the other by Professor P. Mayet (3). The first of these experiments was un-

(1) *Hösten i Danmark i Aaret* 1902. (Statistiske Meddelelser Series 4, No. 13).

(2) « Errechnung des Viehstandes vom 1. Dezember 1897, 1900 und 1902 für einem Kreis jedes Regierungsbezirkes des preussischen Staates sowie fur sämtliche Kreise des Regierungsbezirkes Bromberg, ausgeführt auf Grund der Zählungsergebnisse dieses Jahr fur je ein Zehntel sämtlicher Gemeinden vermöge der Verhaltniszahlen, welche aus den Ergebnissen fur 1892 berechnet sind ».

(3) Dr. P. MAYET: *Stichproben-Erhebungen in der Zwischenzeit zwischen grossen Vollzählungen längeren Periodizität.* (Bulletin de l'Institut International de Statistique. Vol. XIV, No. 2. Berlin, 1904).

successful, while the latter gave particularly good results. As to the causes we would refer to the minutes of the preceedings of the meeting in Berlin in Vol. XIV of the Institute's Bulletin, which also contains a very complete account of the method used by Mayet.

A third experiment of the same kind was made a few years ago by Dr. Johannes Lucht in the Prussian Bureau of Statistics.

This experiment comprised all the provinces of Prussia, Berlin and the Hohenzollern Lands alone being excluded. The idea of the experiment was to investigate with what degree of certainty one could, by means of a representative counting of the stock of cattle in one-fourth of the administrative districts (« Kreise »), calculate the changes from 1920 to 1921 with regard to the number of the various kinds of domestic animals. The calculations were made partly for each of the eleven provinces and partly for the whole country as one.

Of the 528 districts one-tenth was selected four times, there thus being 4 samples of 53, 53, 53 and 52 districts, and each of these four samples was examined separately. The samples were selected as follows : —

The 11 provinces were arranged in the order usually adopted in statistics, and in each province the districts were arranged according to the initials of the names of the districts. Consecutive numbers were then given to the districts in this arrangement. Then were selected all the districts whose numbers ended with 3 (i. e. 3, 13, 23, etc.) and the sample thus arrived at, of 53 districts, was named « Schätzung I ». Thereafter were selected the districts whose numbers ended with 5 (« Schätzung II »), with 8 (« Schätzung III ») and with 0 (« Schätzung IV »).

After the investigation into these 4 original samples had been completed, it was desired to ascertain what influence the relative size of the sample had upon the accuracy of the result, and therefore two new samples were formed by adding together I and III (« Schätzung V ») and II and IV (« Schätzung VI »). Finally, all four tenths were amalgamated into one single sample (« Schätzung VII »).

In all these seven experiments the number of horses, cattle, sheep, goats and pigs were calculated. As regards Sample I there was also calculated the distribution of domestic animals in age groups (« Schätzung VIII »).

While the original Samples I-IV were selected on a mechanical principle (the last figure of the consecutive number), an experiment was also made where the sample was formed by deliberate selection of the districts (« Schätzung IX »). In selecting the districts for this investigation care was taken that the geographic representation was

fairly equal, while the size of the stock of cattle, the prevailing method of feeding and the distribution of the farms according to size were also taken into consideration.

Dr. Lucht has treated the results of these experiments in a very thorough manner in a treatise in the Journal of the Prussian Bureau of Statistics (1). The figures arrived at for the various provinces often differ greatly from the actual figures; for the State as a whole the differences are much smaller, but the author describes the results of the experiments as unsatisfactory on the whole.

IV. — Purposive Selection of Groups.

We shall now describe a number of representative investigations which differ from those referred to in the foregoing in that the sample was not drawn by random selection (neither of units nor of groups), but by purposive selection of such unit-groups as could be presumed to have approximately the same properties as the whole they were to represent.

It sometimes happens that the official statistical Department, after having set up a large and costly apparatus in order to secure statistical information for administrative or legislative purposes, is faced by the situation that information is required regarding other points than those included in the enquiry made. In such cases it will often be unreasonable to make the investigation over again, quite apart from the fact that there may not be sufficient time to do so. It will then often be possible to supplement the results of the totality investigation by help of a representative supplementary enquiry. The following is an example of this.

When the question of the franchise to the Second Chamber of the Hungarian Parliament was to the fore in 1904, a comprehensive statistical enquiry was made by the Hungarian Bureau of Statistics regarding all the male inhabitants of more than 20 years of age. After this large material had been worked up, there was a desire to have these statistics supplemented with information as to certain circumstances which had not previously been mentioned (how long each person had been residing at his actual domicile, and how long he had been occupied in the trade in which he actually worked). As there could hardly be any question of making an entirely new

(1) Dr. Johannes Lucht: *Die representative Methode in der Statistik.* (Zeitschrift des Preussischen statistischen Landesamts, 1922).

investigation regarding all men over 20 years, the Bureau simply made a supplementary enquiry in a number of administrative districts which altogether represented about 5 per cent of the population of the country. The administrative districts which were investigated were so selected that they could be presumed to represent as closely as possible the whole population as regards demography and occupation. On the question forms which were employed in this representative investigation there were — apart from the new questions referred to above — included some of the questions to which answers had already been given in the original, complete enquiry, and it was thus possible to gain a clear conception of the degree to which the sample chosen was really representative of the whole with regard to the data which were of importance to the object of the enquiry. This experiment is said to have given a particularly satisfactory issue, but it is not possible to give any details, as no report on the investigation has ever been published (1).

In the years 1891-1899 the Central Commission of Statistics in Holland made different enquiries into the degree of wealth and birth and infantile mortality. To this end a few towns and a certain number of country places were selected. The selected communes were meant to be representative for the whole country (2).

At a time when the Polish census material of 1921 had not yet been investigated, some preliminary information was required with regard to the size of certain age classes of the population (children of school age and men of military age). A partial investigation was therefore made of the census material, certain districts being selected in each province, and a calculation being made of the ultimate importance of each of the districts investigated when included in the average. The results of this investigation have not been published (3).

We may also mention that a similar investigation to that just referred to, but for a limited area (a single department) was made by the late Polish statistician and astronomer R. Merecki (4).

After having referred in the foregoing to some examples of the application of the purposive method in regard to demographic invest-

---

(1) Information kindly furnished by Dr. Ladislaus von Buday.

(2) *Uitkomsten van een onderzoek naar het verband tusschen welstand, natalitеit en kindersterfte in Amsterdam, Dortrecht, Rotterdam,* en 40 plattenlandsgemeenten. Maandcijfers, 1897, No. 8 ; 1898, No. 10 ; 1899, « Nieuwe Volgreeks» No. 2 ; 1900, No. 4.

(3) Information kindly furnished by Dr. J. Buzek.

(4) Revue Mensuelle de Statistique, Vol. VI, pp. 81-84, Warsawa, 1923.

igations, we will now proceed to deal with another group of representative enquiries on the basis of purposive selection, viz. those for the purpose of providing information as to the conditions under which the population lives and works.

It is true of a large number of such investigations that they are not of much interest in a methodological sense, because, either of necessity or otherwise, controlling of the representative character of the material for investigation has been waived. Thus, most enquiries regarding wages and working hours, etc. are based upon information from which no generally applicable conclusions can in reality be drawn with certainty. These are most often partial investigations without any real representative character.

Naturally there are exceptions to this rule. Thus there is at any rate great probability that the great German wage statistics of 1920 gave in its essentials a correct picture of the height of wages in the various trades in Germany at that point of time. This investigation comprised about 11,700 establishments, with 1,560,000 workmen and 227,000 employees. When selecting the establishments to be investigated importance was attached to both large, medium sized, and small establishments being equally represented. The selection was made by the so called "« Reichs-Arbeitsgemeinschaften » on the suggestion of employers and workmen, and replies to the questions asked were made compulsory (1).

A certain representative character may perhaps also be granted to the great investigation into wages and hours of work in the French industrial establishments in 1891-93. This investigation, which was undertaken by the Office du Travail, comprised about 3,000 establishments belonging to the large and medium sized industries. These 3,000 establishments, which employed a fourth of all the work-people in the large and medium-sized industries, were distributed over all the districts in the country. In the choice of the establishments it was endeavoured to make the sample representative both as regards locality and trade (2).

The investigations made by A. N. Kiaer in the nineties, which formed the starting point for the discussion on the representative method at a number of meetings of the International Institute of Statistics, were representative in the truest sense of the word.

---

(1) *Lohn-und Gehaltserhebung vom Februar* 1920. (Statistik des Deutschen Reichs, Vol. 293, Berlin 1921).

(2) OFFICE DU TRAVAIL : *Salaires et durée du travail dans l'industrie française*. 4 volumes and 1 album of diagrams. Paris, 1893-97.

For use in the preparation of an invalidity insurance scheme in one of these enquiries a number of questions were directed to a large number of persons, selected according to rules which were framed to ensure that the sample represented the whole population. We shall, however, not go into the method used more closely, it being described in the Bulletin of the Institute (1).

The representative method was applied in an investigation made in 1917 in Sweden into the consumption of spirits.

The peculiar system is prevalent in Sweden that the consumption of spirituous liquor is rationed. When a purchase of such beverages is made, the quantity bought and the amount paid for it are entered in the so-called « Motbok », an account book specially designed for this purpose and issued to all who, according to the regulations, are entitled to buy spirituous liquor.

The aim of the investigation was to show the extent of the consumption of spirits by persons belonging to various trade groups and income-classes. The material for the investigation was procured in this way: a number of the companies, who have a retail licence to sell spirits, took copies of the account-books of their customers. There were in all about 600,000 such books, of which something more than 45,000, or about 8 per cent, were included in the investigation. In order to obtain as representative a sample as possible, a number of selling companies, distributed over every part of the country, was first selected, and after these companies had been requested to assist in the investigation, detailed instructions were drawn up regarding the procedure they were to follow. As to the companies which comprised the largest towns, it was decided that the whole of the customers of certain indicated retailing shops should be included. For the medium sized companies, a number of particular districts was selected. In selecting the retailing places and the districts, endeavours were made to obtain as far as possible equal representation of well-to-do and poor quarters, industrial and farming districts, etc. As to the smallest companies, the whole of their customers was included.

No test of the representative character of the sample used was made, and as a matter of fact this would, especially in view of the peculiar nature of the investigation, seem to be impracticable (2).

---

(1) A. N. KIAER: *Obsérvations et expériences concernant des dénombrements représentatifs.* (Bulletin de l'Institut International de Statistique. Vol. IX, No. 2, p. 176).

(2) *Förbrukningen af utminuteringsvis sält brännvin efter yrkesgrupper och inkomstklasser.* Stockholm, 1919.

The method recommended by A. N. Kiaer in the nineties was, by the way, previously used in enquiries of various kinds, as for instance in an enquiry regarding housing conditions and rent undertaken by the Municipal Office of Statistics of Copenhagen in 1885.

The sample used consisted of 36 streets, spread over every quarter of the town. In these 36 streets there were altogether 9,366 dwellings (flats), which was one-seventh of all the dwellings in the town, and the number of inhabitants was 39,350, or similarly one-seventh of the whole population of the town.

The representative character of the sample will also be seen from the following distribution figures (1):

DWELLINGS WITH	36 streets %	Whole town %
One room	14.3	15.0
2 rooms	36.7	36.3
3 »	17.3	16.8
4 »	12.3	12.8
5 »	6.9	7.0
6 »	5.6	4.6
7 »	2.6	2.8
8 »	4.3	4.7
	100.0	100.0

In the investigation which has just been referred to the sample consisted of *all* the dwellings in certain streets, distributed over the whole of the town. A rather modified proceeding has been followed in an investigation into the height of house rent in 1924 in 28 towns in Switzerland. According to the plan of this investigation, information was to be procured as to the house rent for about ten per cent of all dwellings with 2 to 4 rooms. The selection was made in this way: certain streets were selected beforehand, and in these certain house numbers. When selecting the streets an endeavour was made to secure an equal representation of all quarters in the town, and in selecting the houses it was tried, as far as was practicable, to get both old and new houses represented in proportions which corresponded with the actual position. As far as is known, however, no attempt was made to test whether the sample really was representative or not (2).

(1) Tabelvaerk til Kjöbenhavns Statistik, No. 9 (Kbh. 1888).

(2) *Résultats de l'enquête sur les loyers, effectuée au mois de janvier* 1924. (Information de Statistique Sociale. *Year II, vol.* 2. Berne, 1924).

In Norway, an investigation was made, by the application of the representative method, into the distribution of the population according to income, fortune and occupation. The investigation was made on the basis of the tax assessment for 1911, compared with the census of 1 December, 1910.

As to the rural districts, the investigation comprised a sample consisting of about one-third of the circuits in each county. These were selected so that the distribution of the population according to livelihood was as near as possible the same in the selected circuits and in the whole country.

As to the large towns, the random method was used, each fourth or each second house being taken ; in the small towns the investigation was complete (1).

To illustrate how far one can go in the direction of accuracy by applying the purposive method in statistics regarding the distribution of the population according to income, we will describe the following experiment, which was made in Denmark.

Based upon the assessment for income and property tax, a table is worked out every year in Denmark, showing for each assessment district the number of tax-payers and their total income and property. But besides this there is worked out for the whole country a table showing the distribution of the tax-payers according to size of income. As this involves a considerable amount of work, the question has been considered whether it would not be possible to arrive at a sufficiently accurate result for the country as a whole by means of an investigation of a representative sample, comprising one-fifth of the assessment districts. In this experimental enquiry the towns were disregarded.

In the first experiment, 15 of the 76 assessment districts were extracted by mechanical counting out. The sample obtained in this manner is characterised by the following qualities :

Sample *A*	Percentage of Totality
Tax-payers	19.7
Income	20.9
Property	20.9

As these criterions seemed to indicate that the sample was sufficiently representative, an investigation was made for the 15 districts into the distribution of tax-payers according to size of income.

---

(1) *Indtaegts- og Formuesforhold efter Skatteligningen* 1911. (Révénus et fortunes selon l'imposition de l'exercice 1911). Norges officielle Statistik, Series VI, No. 24.

The result is shown in the table below, Col 2. For purposes of comparison the corresponding percentages for all 76 assessment districts are shown in Column 1.

As will be observed, there is very good agreement between the results of the complete and the partial investigations, although both the lowest and the highest income groups are rather too strongly represented in the sample.

Percentage Distribution of Tax-payers.

INCOME GROUPS	1	2	3
	Totality (1)	Sample A (2)	Sample B (3)
800 – 1,000 Kr.	4.58	5.32	4.76
1,000 – 1,200 »	7.38	7.84	7.43
1,200 – 1,400 »	8.21	8.51	8.10
1,400 – 1,600 »	10.22	10.57	10.71
1,600 – 1,800 »	7.56	6.59	7.48
1,800 – 2,000 »	6.54	6.08	6.57
2,000 – 2,500 »	12.56	12.26	12.15
2,500 – 3,000 »	9.09	8.89	8.89
3.000 – 3,500 »	7.80	7.82	7.66
3,500 – 4,000 »	5.84	5.99	5.84
4,000 – 4,500 »	4.58	4.67	4.54
4,500 – 5,000 »	2.91	3.01	2.86
5,000 – 5,500 »	2.50	2.43	2.60
5,500 – 6,000 »	1.60	1.54	1.58
6,000 – 7,000 »	2.63	2.59	2.63
7,000 – 8,000 »	1.68	1.62	1.72
8,000 – 9,000 »	1.13	1.05	1.16
9,000 – 10,000 »	0.75	0.71	0.82
10,000 – 15,000 »	1.53	1.49	1.59
15,000 – 20,000 »	0.43	0.41	0.44
20,000 – 30,000 »	0.26	0.31	0.28
30,000 & over.	0.22	0.30	0.19
	100.00	100.00	100.00

(1) 76 assesment districts, 483, 356 taxpayers.
(2) 15 assesment districts selected at random, 89, 716 taxpayers.
(3) 15 assesment districts selected under consideration of the average income, 97, 086 taxpayers.

The figures above show that the population covered by the sam ple is, both as regards income and property, rather better situated than the totality, as the following figures also show :

Average per tax-payer	Sample A	Totality
Income	3,213 Kr.	3,067 Kr.
Property	14,120 »	13,500 »

It must therefore be probable that a more representative sample could be obtained by taking advantage of the knowledge one has of the average income per tax-payer in each of the assessment districts. This average varies between 1.744 Kr. in the poorest assessment districts and 5,392 Kr. in the richest. Now in order to exploit this factor, the 76 assessment districts were arranged according to average income, and thereafter the assessment discricts thus arranged were divided into three groups, each group comprising as nearly as possible one-third of all the tax-payers. For each of these groups the average income per tax-payer was calculated, and thereafter in each of the groups 5 assessment districts were taken, selected in such a way that (1) the number of tax-payers in the assessment districts selected formed approximately one-fifth of the whole number of tax-payers in the group, and (2) the average income per tax-payer in the selected districts as nearly as possible corresponded to the average income per tax-payer in the whole group. The sample thus obtained was characterised by the following percentages :

Sample *B*	Percentage of Totality
Tax-payers	20.1
Income	19.9
Property	20.3

The table below gives more details :

	Number of Tax-payers		Average Income per Tax-payer	
	Totality	Sample *B*	Totality	Sample *B*
			Kr.	Kr.
Group I	162,615	32,619	2,415	2,359
Group II	160,668	32,327	3,076	3,078
Group III	160,073	32,140	3,720	3,693
	483,356	97,086	3,067	3,040

This new sample having been formed, the distribution of tax-payers according to income was examined as with the first sample. The result will be found in Col. 3 in the table. It will be seen that the agreement with the percentage arrived at by the complete investigation (Col. 1) is almost ideal.

In order to show how little the difference between the results of the complete and the partial investigation means in relation to the differences from year to year, a summary of the distribution of income in 1919 and 1921 is given in the table below, as well as for 1920, partly according to the complete and partly according to the partial investigation.

	1919	1920			1921
	Totality	Totality	Sample *A*	Sample *B*	Totality
800– 1,400 Kr.	24.13	20.17	21.67	20.29	24.48
1,400– 3,000 »	47.59	45.97	44.39	45.80	46.64
3,000– 5,000 »	17.60	21.13	21.49	20.90	18.88
5,000–10,000 »	8.35	10.29	9.94	10.51	8.50
10,000 Kr. & over	2.33	2.44	2.51	2.50	1.50
	100.00	100.00	100.00	100.00	100.00

Partial enquiries have in particular been applied to agricultural statistics in the preparation of « crop forecasts » and the « crop returns » proper. In both respects the procedure is, in principle, the same : as it is impracticable to obtain information from every single farm, the statistics are built up upon a number of reports by farmers, local authorities, agricultural advisers, etc. in all parts of the country, it being presumed that these reports are sufficiently representative to form a basis for results of general validity.

It is obvious that statistics of this kind now and then need checking by means of enquiries founded upon direct measurements. That these may be of great value, even if they are only carried out for a part of the farms, is proved by the example of the extra-ordinary crop statistics compiled in Norway in 1918.

This enquiry covered the weight of the grain yield of the various kinds of grain and peas and the weight of the potatoes, turnips and hay harvested. The investigation was made by local commissioners, and it was laid down that the weighing of the crops should take place

at a certain number of different places in the area of each commissioner. In selecting the farms to be examined importance was attached to making the investigation cover soil of medium yield, and that large, medium sized and small farms were represented. For hay, turnips and potatoes a procedure was adopted which resembles that which is often used in the calculation of the wood content of a forest; right-angled sections of a definite size were marked out and the yield of these was weighed. This check-enquiry thus in more than one respect made use of the representative method (1).

The representative method was used in a peculiar manner in the Swedish agricultural statistics in the years from 1913 to 1921. Every year in this period local agricultural-statistical enquiries were made in the following manner. The country is divided into 25 counties; the area of each of these was divided into 8 districts, one of which was investigated each year, so that the whole country was completely covered in the course of 8 years. For all the farms in the districts examined statistics were compiled, comprising the utilisation of the area, seed sown, crop yield and stock of cattle. In these local enquiries it was thus taken for granted that the eighth part of the country's farms, which were each year subjected to examination, were representative of all the farms with regard to the factors named above. It is thus clear that the value of the whole system depended upon the manner in which the dividing of the counties into investigation districts was made. This division was made once and for all by the Central Bureau of Statistics in consultation with the executive committees of agricultural societies which are to be found in all the counties.

In Section III we have mentioned an investigation where the object was to determine the utilisation of area for all groups of farms together. The representative method has, however, also been employed in more thorough investigations in this respect.

In Czecho-Slovakia, on the occasion of the passing of an Act in 1918 regarding the expropriation of land by the State, the State Bureau of Statistics made an investigation into the utilisation of area on farms of various sizes (2). As for technical reasons it was impossible to carry out this investigation for the whole country, a sample, comprising 12 of the 464 districts into which the country is divided, was examined.

---

(1) Norges officielle Statistik, Series IV, No. 167. Kristiania, 1920.

(2) According to a treatise in the Czech language by Dr. J. Auerhan in the Czecho-Slovakian State of Vestnik vol. 1, books 6-7.

In the 12 districts which formed the sample examined, there were 44,773 farms with a total area of 279,206 hectares and an agricultural area of 126,158 hectares. This latter area was 2.03 per cent of the whole agricultural area of the country.

The 44,773 farms were divided according to the size of their land into 9 groups, by which the following figures were arrived at :

	Number of Farms	Agricultural Area — Hect.
Less than 5 ha.	37,759	35,284
From 5 to 10 ha.	3,280	16,182
» 10 » 20 »	2,366	24,040
» 20 » 50 »	1,120	22,031
» 50 » 100 »	138	4,976
» 100 » 200 »	47	2,834
» 200 » 500 »	35	3,212
» 500 » 1000 »	8	1,382
» 1000 and more	20	16,217
Total	44,773	126,158

For each of these 9 groups an investigation was now made as to the distribution of the agricultural area according to the cultivation of wheat, rye, barley, oats, potatoes, sugarbeet, forage crops and other crops, and fallow land.

The detailed results need not be given here. It shall merely be observed that a rather considerable difference was ascertained between the various groups, the cultivation of wheat, barley and sugar-beet specially being proportionately more extensive on the big farms than on the small ones, whilst on the other hand the smallest farms had comparatively large areas under rye and potatoes. The maximum for wheat, barley and sugar-beet was not, however, found in the group of the biggest farms, but in the group 100-200 hectares, and Dr. Auerhan attached a certain amount of importance to this fact having regard to the circumstance that the Act mentioned spares just this group.

An investigation of the same kind as that made by Dr. Auerhan in Czecho-Slovakia regarding the utilisation of the area in farms of various sizes has been made in Denmark as an experiment.

Of the 90 parishes in Prästö County, 20 parishes were first selected by mechanical counting out. The geographic distribution of the parishes thus arrived at was not particularly even ; some parts of the county were strongly represented in the sample, others not at all.

Too much weight was not, however, attached to this, as it was of course possible that the sample despite the somewhat uneven geographic distribution might be representative. How far this was the case was investigated by the help of the following criterions.

In percentage of the totality (the County) the figures for the sample (the 20 parishes) were :

	Percentage
Area	16.8
Hartkorn (1)	18.9
Horses	20.7
Cattle	21.5
Of these, milch cows	21.6
Pigs	21.9

The circumstance that the percentage for the area is less than that for the « hartkorn » shows that the sample as a whole covers soil of rather greater fertility than the County as an average. This also harmonises with the fact that the stock of cattle in the 20 parishes, as the figures show, is comparatively greater than in the County as a whole.

As it thus proved that the mechanical selection did not give a sample which, in regard to representative characteristics, satisfied the strict demands which must be made in an investigation of this kind, the mechanical selection was dropped and a new sample was formed in the following manner.

As the manner in which the agricultural area is utilised harmonises both with the fertility of the soil and the stock of cattle, these two factors were used as a guide in the selection of 20 parishes : fertility (that is to say proportion between area and « hartkorn ») as the primary, and the stock of cattle per area unit as the secondary criterion.

For each of the 90 parishes in the County the proportion between agricultural area and « hartkorn » was calculated. Thereafter the parishes were arranged according to these ratios and divided into four groups, each comprising one-fourth of the total agricultural area in the County. In each of the four groups the parishes

(1) The « Hartkorn » is a quality measure for the soil.

were arranged according to the proportion between the number of heads of cattle and the agricultural area, and thereafter, having regard to this ratio in each group, 5 parishes were selected comprising as nearly as possible one-fifth of the area.

The 20 parishes which were selected in this manner covered 2,097 of the County's 10,207 farm properties, and 28,488 of its 139,698 hectares of agricultural area.

In the sample, and in each of its four groups, the following proportions were found:

	Number of ha per « Tønde Hartkorn »	Per 1000 ha. Agricultural Area			
		Horses	Cattle	Milch-Cows	Pigs
Group I.	4.54	203	918	525	957
» II	5.09	206	834	479	714
» III	5.52	197	804	473	658
» IV	6.33	184	782	460	754
Whole sample, gr. I-IV . .	5.29	198	835	486	772
Whole County	5.29	196	835	490	745

As will be seen, it had thus been possible to select in this manner 20 parishes in which the average quality of the soil and the relative stock of cattle were exactly as the average for the County as a whole.

To what degree the new sample (Sample 2) surpassed the mechanically selected one (Sample 1) with regard to the representative characteristics will appear from the following figures:

	Percentage of the Totality	
	Sample 1	Sample 2
Agricultural area.	16.8	20.3
« Hartkorn »	18.9	20.2
Horses.	28.7	20.4
Cattle	21.5	20.3
Milch cows.	21.6	20.0
Pigs	21.9	21.0

With regard to the summary distribution of the agricultural area according to its utilisation the following percentages may be given :

	Totality (County) %	Sample 2 %
Grain to maturity:		
Wheat	4.5	4.1
Rye	6.2	6.8
Barley	14.0	14.4
Oats	10.0	10.0
Mixed seed	8.3	8.5
Buck-wheat, pulse	1.2	0.8
	44.2	44.6
Root crops	12.8	13.0
Seeds, etc.	3.3	3.4
Fallow	7.7	8.0
Green fodder and grazing areas	28.3	27.2
Gardens, etc.	3.7	3.8
	100.0	100.0

The experiment of Professor *Mayet* as to applying the representative method in enquiries regarding stocks of cattle (Section III) had attracted so much attention that the Austrian Central Commission of Statistics decided to try the representative method on the material at hand for the Duchy of Salzburg and for Bohemia with regard to the stock of cattle in the years 1890 and 1900. In the former case the growth from 1890 to 1900 was reckoned on the basis of the actual counts for 25 of the 154 municipalities. As regards Bohemia the investigation was much more comprehensive, as, on the basis of the actual counts for 1,049 of 7,407 municipalities, a converse calculation was made (from 1900 to 1890).

In both cases the degree of accuracy of results arrived at by the representative investigation could be checked, as both for 1890 and 1900 there were complete counts.

While Mayet selected his sample at random, the Austrian experiment was based on purposive selection. When selecting the

sample the area of the enquiry was divided into a number of « natural districts » (4 for Salzburg, 13 for Bohemia), thus ensuring that in the sample there were represented districts with widely different agricultural conditions. In the selection of the municipalities in the « natural districts », the population was taken into consideration ; both large, medium sized and small municipalities were included, but so that the big municipalities were proportionately more strongly represented. The results of these experiments were on the whole not very satisfactory (1).

The representative investigations regarding stocks of cattle which we have referred to in the foregoing (Pfaundler & Weyr, see above ; Kluge, Mayet and Lucht , see Section III), are all experiments made for the purpose of testing the applicability of the representative method in this special sphere. There are, however, examples showing that representative livestock censuses have been made as a substitute for complete censuses.

Such a partial livestock census was, for instance, taken in Sweden in 1915. The sample used in this case consisted of rather more than 100 parishes, spread over all parts of the country and so selected that districts of different character and soil conditions were as a far as possible represented. The results of the census in the selected parishes were compared with similar figures for 1913 and 1914 and, on this basis, the probable increase or decrease of the livestock numbers for the country as a whole was calculated (2).

The representative investigations regarding the livestock referred to in the foregoing have this in common, that the unit operated with was some administrative district or other.

Attempts have, however, also been made to determine the number of livestock by undertaking a count of the animals on a certain fraction of the farms in all districts of the country. This method has repeatedly been used in Norway.

In 1915 the Norwegian Central Bureau of Statistics made a representative livestock census by counting the animals on an average of 15 farms in each district, in all of 11,188 farms, or 6.5 per cent of all independent farms with livestock. On the question forms information was given as to the number of the various kinds of domestic animals (divided into age groups) at the end of September 1914

(1) Dr. Richard Pfaundler and Dr. Franz Weyr: *Die stichprobenweisen Viehschätzungen. Eine kritisch-methodologische Untersuchung.* (Statistische Monatschrift. 1906, pp. 551 & 695).

(2) Statistiska Meddelanden. Ser. A. Vol. I, N° 7. Stockholm, 1915.

and 1915. After the question forms had been filled in and returned to the Central Bureau, a count was made for each of the 18 counties in the country, and a corresponding count *for the same farms* was thereafter made on the basis of the question forms preserved on the files of the Bureau from the last (complete) livestock census, which was taken in 1907. By means of this identification of the farms it was possible to calculate the percentage of change in the stock of animals in the period from 1907 to 1914, and as it was taken for granted that the selected farms in each county were representative of all the farms in the county, it was possible to calculate approximately the size of the stock of animals in each county in 1914 and 1915.

In order to secure as representative a selection of farms as possible, an instruction had been drawn up for the municipal authorities to the following effect: 1. The farms selected were to be of different sizes, small, medium sized and large farms being represented in proportions which might more or less be regarded as corresponding to the general distribution of farms in the district. 2. The farms were to be situated in the various parts of the district; it was emphasised that an equal geographic division was particularly necessary, where the natural conditions of the farms were essentially different in the various parts of the district. 3. In view of the comparison with the figures for 1907 it was necessary to include only farms which in the intermediate period had neither been reduced by sale of land or increased by purchase of land.

As has already been mentioned, the sample formed 6.5 per cent of all farms in 1907. In that year the stock of animals on the farms investigated was in percentage of total stock of animals at that time:

Horses	11.4 %
Cattle	10.2 »
Sheep	8.5 »
Goats	7.0 »
Pigs	9.3 »

It will be seen that the sample has not been fully representative with regard to the composition according to size of farm; the small farms have obviously been too weakly represented in the sample (1).

A similar investigation was made in 1916, this time the counting of animals being made on the same farms as in 1915, and also on a number of small farms which were not included in 1915 (2).

(1) Norges officielle Statistik. Series VI, No. 64 Kristiania, 1916.
(2) Norges officielle Statistik. Series VI, No. 92, Kristiania, 1917.

In 1923 and 1924 representative livestock censuses have been made on a plan similar to that described above, but in such a manner that, for comparison, the figures of a complete census in 1918 have been used (1).

The method described above, where the sample was formed of a certain fraction of the farms in all districts of the country, was used from and including the year 1921 in Finland, not only for the purpose of determining changes in the number of livestock, but also in the annual statistics of the utilisation of the agricultural area.

The procuring of the material for the annual agricultural statistics is entrusted to the local agricultural societies. The executive committees of these societies receive from the Ministry of Agriculture lists of all the farms in each district, grouped according to size. These executive committees attend to the appointment of census agents in every district, and these, on the basis of the lists referred to, select a tenth of the farms in each size-group, endeavours being made in this selection to ensure that the sample is in every respect as representative as possible. The particulars as to the number of livestock and the utilisation of the area, which are procured for the selected farms with the assistance of the agents, are compared with similar information from a complete investigation made in 1920 and, with this as a basis, the figures for the various districts and the country as a whole are calculated (2).

Despite the great improvements which agricultural statistics have undergone in the various countries during the course of years, it must be said that on the whole we have not got much further than giving a summary picture of the apparatus of production of agriculture and the size of the gross production. We will not, however, be able to remain here in the long run, and in several countries attempts have, in fact, been made to go into the question more deeply and to explore the mysteries of farming itself, and even to determine the profitableness under given conditions.

Such investigations, which as a rule are due to private initiative or are undertaken under the auspices of the agricultural organisations, are, however, compelled to work on a very slender foundation, viz., the comparatively few working accounts which can be procured through voluntary channels. It will always be possible to dispute

---

(1) Norges officielle Statistik, Series VII, No. 117, Kristiania 1924.
Norges officielle Statistik, Series VII, No. 153, Oslo, 1925.

(2) Finlands officielle Statistik III. Lanthushallning, 18. Jordbruk och boskapsskötsel ar 1921. Helsingfors, 1924.

whether and to what degree this account-material can be regarded as being representative, and this must naturally to a great extent reduce the usefulness of the results.

However, so far as one can see into the future, there will scarcely be any prospect of it being possible to procure a mass-material consisting of working accounts in such quantity and quality that really representative investigations can be made by this means. The question is, then, whether one cannot approach the same goal by means of an intensive exploitation of the information procurable through simple questions to all farmers.

As the object, however, demands a far-reaching division of the material and numerous combinations of the information, the work may well be insurmountable if all the material is to be included in the statistical treatment. Therefore here we have a domain where it is recommendable to use all the help of the representative method which considerations as to the accuracy of the results permit.

In the following a description will be given of a comprehensive experiment in this direction, made during the past year in the Danish Department of Statistics.

In the agricultural census in 1923 very detailed information was obtained for every single farm as to the area and its utilisation, the stock of cattle, the number of labourers employed on the farm, the farmers participation in different cooperative societies, etc. After a summary count had been made of the various columns for each parish, the investigation was carried further by combining the information from the various farms in different ways. Thus in this second section of the investigation merely a representative part of the material was worked up in order to save time and money.

It was determined to make the size of the representative sample about 20 per cent of the whole, which means that the sample was to form about 20 per cent also in each of the country's 22 counties. In selecting the sample the procedure was as follows:

For each of the 1,809 parishes the number of cows per 100 hectares of farm area was first calculated, and the parishes were arranged according to these ratios in each of the 22 counties. In this order the parishes in each county were divided into groups of 5 parishes, and from each of these groups was selected the parish whose agricultural area most nearly amounted to one-fifth of the total agricultural area of the group. In some cases, where the parishes varied greatly in extent, the material was adapted by adding two parishes together or, conversely, dividing a large parish.

The parishes thus selected were now drawn on a map, in order

to examine whether the various parts of the county were fairly equally represented. Where this was not the case it was sought, by means of exchanging the selected parishes with others where the number of cows per 100 hectares of agricultural area was about the same, to bring about the desired equal geographic representation.

This preliminary sample was now, in each county, tested with regard to the following factors : 1. Number of farms; 2. Total agricultural area ; 3. Grain area ; 4. Rootcrop area ; 5. Grazing area ; 6. Meadow area ; 7. Number of horses ; 8. Number of milk cows ; 9. Total number of cattle ; 10. Number of pigs, it being noted whether the sample, with regard to these various factors, represented approximately one fifth of the whole county. Where considerable differences were found in one or more respects, an examination was made to ascertain from which parish or parishes each of the differences originated, and thereafter an attempt was made to remove the differences, as far as possible, by exchanging these parishes with others where the proportion between the size of the agricultural area and the number of cows was the same. Furthermore, care was taken that the geographic representation was not too disturbed.

Where the figures for a county in some respect or other were especially small, it is natural that the incidental circumstances in the selection of parishes might have a far-reaching effect (this is, for instance, true in regard to some counties with respect to the meadow area). In such cases the demands as to the representative character of the sample had naturally to be somewhat modified.

To what degree it was possible to compile a representative sample will appear from the table below : —

	Percentages of Totality		
	Sample for the Islands	Sample for Jutland	Total Sample for Denmark
Number of farms	20.8	20.7	20.8
Total agricultural area	20.6	20.9	20.8
Grain area	21.0	20.7	20.8
Root-crop area	21.0	20.8	20.9
Grazing area	21.1	20.8	20.9
Meadow area	19.5	20.9	20.7
Number of horses	21.0	20.6	20.8
Number of milk cows	20.8	20.8	20.8
Total Number of cattle	20.8	20.9	20.9
Number of pigs	20.8	21.1	21.0

V. — Purposive Selection of Groups Combined with Random Selection of Units.

The investigations we have referred to in the foregoing have all been classifiable either under the random method or the purposive method. It is clear, however, that it is unnecessary always to keep to the one or the other of these methods. If one has a large and varied sphere of investigation, where the application of the random method is impossible for practical reasons, it will often be recommendable to combine both methods in selecting the sample.

There is an example of this in the investigation into the distribution of income in the Norwegian population which was once made by A. N. Kiaer.

First of all a number of towns and rural districts, spread over the whole country, was selected. In each of the districts chosen, men of an age of 17, 22, 27, 32 etc. were made the subject of enquiry and, among these, were finally selected those whose names began with certain initials. The sample, which formed 3.3 per cent of the male population in the towns and 1.6 per cent of the rural population, proved after various tests to be very satisfactory (1).

Another example of a combination of the purposive and the random selection is a study of the incomes of persons over 65 years old, recently undertaken in Massachusetts. The Commission on Pensions, which conducted the investigation, decided that a full enumeration was impracticable, and that no more than 20,000 out of 200,000 persons over 65 years could be interviewed. A house-to-house canvass was necessary. A complete use of the random method was impossible. It was decided, therefore, to select the districts to be sampled on a representative basis, so as to secure proportional representation of rural and urban, and manufacturing and commercial communities, etc. But within each community reliance was put upon the random principle. Generally, every third or fourth person on an alphabetical list was interviewed (2).

---

(1) A. N. KIAER: *Die repräsentative Untersuchungsmethode.* (Allgemeines statistisches Archiv., 1899).

(2) Information kindly furnished by Professor ALLYN A. YOUNG.

VI. — Incidentally taken Samples, appearing to be Representative.

In the investigations which are referred to in the foregoing sections, very different modes of procedure have been applied, but all these investigations have this feature in common, that the directly investigated portion (the sample) has been the result of a more or less free choice. This freedom in the selection of the sample, however is not always possible. There are cases where one is bound to use a material which may incidentally be available or can be procured. If one is quite unable to judge whether such an incidental sample has approximately the same properties as the whole, the investigation naturally cannot be called representative. But if one has the means of testing the representative character of an incidentally taken sample, and if the test gives a satisfactory answer, the investigation in question can, of course, replace a complete investigation in the same degree as if the sample had been selected at random or by purposive selection.

Incidentally taken samples are, for instance, used in most investigations regarding the consumption of foodstuffs. Investigations of this kind are, as a rule, made according to methods which only in a figurative sense can be described as being representative. Sometimes the Enquiry Method, and sometimes the Monographic Method is used, but it is seldom that the possibility offers of testing to what degree the results may be presumed to be generally applicable. In some cases, however, such a test has been tried.

We will describe one or two examples of this.

In 1915 a fairly comprehensive investigation was made in Munich into the consumption of the most important kinds of foodstuffs by the population. From about 5,000 households information was secured voluntarily as to the quantity of the actual consumption of the commodities in question in the period from February 1st to 26th. Even if in the procuring of this material various measures were taken in order that the various classes of society might be represented, one circumstance alone — the fact that replying to the question forms sent out was voluntary — naturally gave the sample a certain contingent character. And so in order to obtain a means of judging whether the sample was sufficiently representative for the purpose of the investigation, the 5000 households were divided according to size, and the figures thus obtained were compared with

the corresponding distribution figures in the general census of population of 1910. By this means the following figures were found, which show the number of households in the sample in percentage of the total number of households in each size group. (1)

Households of	1	person	0,80 %
Id.	2	persons	2.98 »
Id.	3	id.	3.85 »
Id.	4	id.	3.78 »
Id.	5	id.	3.34 »
Id.	6	id.	3.37 »
Id.	7	id.	3.31 »
Id.	8	id.	3.07 »
Id.	9	id.	2.33 »
Id.	10	id.	4.10 »
Households of more than 10 persons			3.20 »

Another example, which is more interesting, in a methodological sense is the investigation into the consumption of bread and flour which was made in the town of Mannheim in the winter of 1915.

In this investigation there was at hand information concerning the consumption in 416 families totalling 2,164 persons, or only 1 % of the population of the town. The 416 families comprised members of the Town Council, employees of the city corporation, and a number of working-class families which the trades unions had induced to give the required information voluntarily. This « sample » was thus neither formed by random nor purposive selection; it was a material which one incidentally had at disposal. It was therefore of importance to find a means of testing whether this incidental material more or less reflected the whole. The following procedure was adopted: —

For each of the 416 families there was information as to the number of persons and the number of rooms in the house, so that one could calculate the fraction which indicated the « density of habitation » of each family. As one already had information as to the distribution of the whole population according to « density of habitation », and as the density of habitation can be regarded as a very good guide to the greater or lesser prosperity of the family (a factor of great importance in the consumption of bread and flour), this factor was used to test the representative character of the sam-

(1) Dr. Georg Welker: *Die Münchener Erhebung über den Lebensmittelverbrauch* in Februar 1915. Munich, Berlin & Leipzig, 1916.

ple. The result of this test is shown by the following figures, where the group I — II means the lowest density of habitation, group VI the greatest.

DENSITY OF HABITATION	Percentage Distribution of Families	
	Whole Population	The Sample
Group I–II	11.6	16.1
» III	23.3	27.2
» IV	34.3	32.2
» V	22.5	15.7
» VI	8.3	8.8
	100.0	100.0

It thus appears that the sample was fairly representative with regard to distribution according to degree of prosperity, even if the most prosperous families (with the lowest density of habitation) was somewhat more numerously represented in the sample than in the population as a whole. (1)

VII. — Summary.

In the foregoing we have referred, to a more or less detailed extent, to 50 statistical investigations with the application of the representative method. In selecting these 50 investigations we have allowed ourselves to be guided by the desire to make our collection of examples as instructive as possible ; it has not, however, been possible to avoid a certain casualness, because the selection of the examples, has after all been governed by the authors' knowledge of the literature on the subject, and this must, despite valuable assistance by a number of colleagues, be admitted to be both defective and spasmodic. The consequence of this is that we can hardly regard the 50 investigations as a « representative sample » in its proper sense, although this is not saying that an analysis of our collection of examples cannot give hints which are of interest.

(1) Dr. Sigmund Schott: *Der Mehlverbrauch ausgewählter Haushaltungen in Mannheim.* (Beiträge zur Statistik der Stadt Mannheim. 3 Sondernummer). Mannheim, 1915.

If we distribute the 50 investigations according to the years in which they were made, we will find a numerical series which gives us the following main impressions : —

Prior to the middle of the nineties of last century the representative method was rarely used in practical statistics, but in the decade round about the turn of the century a not inconsiderable interest was being taken in the method ; it was in this period that the question was brought up in the International Statistical Institute, where it was discussed in the sessions of 1895, 1897, 1901 and 1903. The investigations made in that period had, characteristically enough, to a great extent the character of experiments. — In the following decade, however, comparatively few representative investigations seem to have been made ; this may perhaps be incidental as to cause, but it would seem to be not unreasonable to connect it with the scepticism towards the method which had been voiced at the meetings of the Institute and the rather reserved recommendation which was finally given to the method in 1903. — After the outbreak of the great war the number of representative investigations suddenly and rapidly increased — perhaps as an expression of the fact that the extraordinary demands of that time upon the statistical service compelled statisticians to utilise a means which, under more normal circumstances, was less favoured. — The numerical series culminates in the last quinquennium, which is doubtless not unconnected with the circumstance that public finances are everywhere strained, while at the same time there is a sustained demand upon official statistics for greater service and higher speed.

Whilst we may take it for granted with a fairly calm conscience that our examples are, at any rate, an approximately « representative sample » as regards distribution according to date, the position is otherwise when we examine the distribution according to countries. We can hardly expect that our collection of examples should give a correct idea of the geographic extension of the representative method. It is probable, for instance, that the fact that the three Scandinavian countries appear with strikingly numerous examples is to some degree connected with the circumstance that the author is a Dane.

The large number from the Scandinavian countries are, however, scarcely wholly misleading. That the representative method has been utilised to an especial degree in Norway is naturally explained by the fact that the Norwegian A. N. KIAER's energetic attempts and warm defence of the method twenty or thirty years ago have acted as a sort of quiet propaganda even up to the present

day. Neither is it to be wondered at that this influence, which naturally has been strongest in Kiaer's own fatherland, has to some degree extended to Norway's neighbours, Denmark and Sweden. Outside Scandinavia, the method seems to have been utilised especially in Germany and England; hitherto it has attracted least interest in the Latin countries, and this despite the fact that France is no doubt the country in which the representative method, in an undeveloped form, was earliest applied. It may further be added that our examples have been gathered from no less than 15 different countries, while isolated examples of the application of the representative method may be found everywhere where statistics have arrived at any particular stage of development in a methodical sense (1).

If we next ask in which domain of statistical research the representative method has been applied to any marked degree, our collection of examples displays an almost equal distribution between the three great groups: 1° Demographic Statistics; 2° Agricultural Statistics and 3° Social Statistics. Outside these three groups there are comparatively few representative investigations, but examples can also be given to show that the representative method has shown the way to the exploration of the important, but for statistics not easily accessible, domains which lie on the boundary between economics and social policy (see, for instance, *Sigmund Schott's* investigations regarding landed properties in Mannheim, mentioned in Section II). Whilst the demographic and social statistical investigations which have been made by the application

---

(1) It must here be admitted that the incompleteness of the foregoing collection of examples is especially prominent on two points, viz. with regard to Russia and North America. That we have not included the Russian representative investigations, which particularly were in respect of national economic and social conditions, is due to the fact that the information as to the methods used are easily accessible in AL. KAUFMANN's instructive work « Theorie und Methoden der Statistik » (p. 362-374). As to America, it must doubtless be admitted that the official statistics in both the United States and Canada have made wide use of partial investigations as substitutes for complete statistics, but in the opinion of the author the methods used there are on the whole not of such a kind that they can be termed « representative » in the narrower sense which this expression, in our opinion, really ought to symbolise (see Section 1). This applies to the numerous partial investigations which have been made by the United States Department of Labor and by Canadas' Dominion Bureau of Statistics, and it also applies, for instance, to the interesting efforts for the promotion of the economic use of Canadas' natural resources made under the leadership of the so-called « Commission of Conservation ».

of the representative method have often been very far-reaching, this cannot be said to the same degree of the investigations made with purely economic phenomena as their object. Thus, the representative investigations into agriculture as a rule only concern such outer and more superficial matters as the utilisation of the area, live stock, etc. ; in one or two cases, however, the method has been used with success in investigations of an economic kind, where it has indeed been the desire to go more deeply than usual (see, for instance, the Danish agricultural statistics of 1923, referred to in Section IV).

In the analysis of our collection of examples there is still lacking mention of the methodical side of the matter. Purely as a matter of form we have, in the foregoing section arranged the 50 examples on the principle used in the selection of the sample. The main interest is not, however, attached to whether the sample is formed by random selection of units, by random selection of groups, or by purposive selection ; as a rule it will rather be purely concrete circumstances than consideration of principle that are of most importance in this respect. Whether the investigation will be successful does not, as a rule, depend so much upon the nature of the method used as upon the details of the manner in which the method is applied. A contribution with the object of ventilating this aspect of the matter will be given in the following section under the title of « Criticism ».

VIII. — Criticism.

The end of the resolution adopted at the Session in 1903 of the International Institute of Statistics in regard to the Representative Method reads : « ...afin qu' un rapport soit présenté sur les applications nouvelles de la méthode et sur la valeur des résultats statistiques obtenus ».

We have attempted to fulfil the first part of this task by referring to a number of investigations which have been made by the application of the representative method. On now passing to the other part of the programme (« la valeur »), we consider it to be unnecessary to embark upon a closer description of those investigations where the method has indisputably proved to be applicable. Such cases speak for themselves.

But on the other hand, there have been a number of experiments and enquiries which have not been successful, and therefore more or

less serious doubts have been raised as to the practicability of the method in general or in special cases where it has been tested. If this pessimism which has thus found expression is warranted, it will be a very weighty argument against the method, and therefore it is of the greatest importance to find out the reasons why a number of representative enquiries have resulted unsatisfactorily. The contents of this section will therefore mostly be in the nature of a criticism of this criticism.

It must first be pointed out that one kind of unwarranted pessimism is founded in the fact that greater demands are made as to the accuracy of the results than is justified if the concrete circumstances of the kind of enquiry or the material used are taken into consideration.

As an example it may be mentioned that in the enquiry which was made in Bohemia by Pfaundler and Weyr (see section IV), the idea was to determine the changes in the stock of animals in a ten year period, during which there had, at any rate as regards some of the animals, been very great changes. (The number of sheep had, for instance, been increased by 86 per cent, the number of pigs decreased by 25 per cent). It is obvious that in such a case one could not expect any great degree of precision in results which were obtained by a partial investigation (1).

In many cases where a representative enquiry really must be acknowledged as having turned out unsatisfactorily on some points or as a whole, it will be seen on investigating the question more closely that the unsuccessful issue is not due to the method as such, but to the manner in which it has been applied in the particular case.

Doubtless the error most frequently committed in representative enquiries is that a sufficiently representative sample is not secured. It must be admitted that this presents an exceedingly difficult problem, but so much the more necessary is it that one should seek to learn a lesson from the errors which have been committed in the past.

It must be remembered that experience shows that even if the sample is selected according to an absolutely « neutral » criterion, a « coloured » picture may be obtained. In the town of Mannheim, for example, statistics were compiled regarding the number of children in each family. The sample comprised the families whose

(1) It is furthermore my opinion that the representative method in the enquiry in question was by no means applied in an irreproacheable manner. This is, however, a point to which I will revert later.

names had the initial letters A, B and M. It turned out, however, that names with these initials were especially numerous among Jewish families, and as the children of the Jewish families were particularly numerous, the enquiry gave a misleading result (1).

There is another similar example in A. N. Kiar's previously mentioned investigation into incomes, where the sample consisted of men of the age of 17, 22, 27 years, etc., these being presumed to present the five-year age classes to which they belonged. This was in so far the case, but the conformity aimed at between the age-distribution of the persons investigated and the age-distribution of the population according to the census was not fully attained, because there is a tendency in population censuses to give « round » age figures, especially in the older age classes.

Many examples could be given of a sample being less representative than expected, because one has rejected or passed over units which, for some reason or other, were not regarded as being suitable for inclusion in the enquiry. In an investigation into economic conditions in Russia, information was obtained from a certain number of farms in each district. If the investigator did not find the farmer at home, he substituted that farm with its nearest neighbour. It turned out that in the sample the more well-to-do farmers had for the most part been included, and thus the sample did not give a proper view of the totality. The reason was that the rejected farms were for the most part owned by peasants who, owing to poverty, had to work as day labourers on other farms, and therefore were not at home when the investigator called at their farms (2).

There was a very similar experience in an enquiry which was made by the British Ministry of Health regarding families attacked by influenza. The houses where the investigator found no one at home were replaced by the next one in the street. This made the sample misleading with regard to the composition of the families, as the « closed » houses were for the most part those occupied by families with no small children (3).

In the previously mentioned enquiry regarding unemployed in England, each of the « tabbed » claims was regarded as making the centre of a small group of claims, each of which should furnish one claimant, and arrangements were made that the first person in any group coming to the Exchange should be interviewed. It has,

---

(1) Dr. Sigmund Schott. Statistik, 3. Aufl. p. 44. (Leipzig-Berlin, 1923).

(2) Al. Kaufmann. Theorie und Methoden der Statistik, p. 365. (Tübingen, 1913).

(3) Journal of the Royal Statistical Society, July 1924, p. 568.

however, rightly been pointed out that this practical arrangement has possibly had the effect that the sample became less representative than if the rule only to include the « tabbed » claims had been strictly observed (1).

In many cases it is easy to prove that the sample used in so-called « representative » enquiries, suffered from grave defects in a representative sense.

In the investigation into the utilisation of area on farms of various sizes which was made in Czecho-Slovakia in 1918, in selection of the sample care was taken that both fertile soil and mountainous districts should be represented, and all parts of Bohemia represented. Dr. Wilhelm Winckler, however, who has subjected this investigation to a critical examination, gives the following figures for comparison of the character of the 12 districts and the whole of Bohemia :

	The 12 Districts	Whole Bohemia
Density of population per sq. km.	96	130
Distribution of population according to occupation, percentages:		
Agriculture	47.30	36.64
Manufacturing and Mining	30.39	36.73
Trade	8.30	10.80
Public services and liberal professions	14.01	15.83
Language distribution, percentages:		
German-speaking	11.05	36.76
Czech-speaking	88.93	63.19
Utilisation of agricultural area, percentages:		
Wheat	15.38	9.83
Rye	16.77	21.69
Barley	11.45	10.05
Oats	9.61	16.36
Potatoes	8.53	9.87
Sugar-beet, etc.	11.74	4.74

(1) Journal of the Royal Statistical Society, July 1924, p. 564.

On the basis of this check-test Dr. Winkler considers it very improbable that the sample selected has been sufficiently representative (1).

Another example of an unfortunate selection of the sample is presented by the already mentioned investigation regarding the changes in the Bohemia stock of animals in the period between 1900 and 1890. The sample comprised 1,049 of the 7,407 municipalities, and the 1,049 municipalities were distributed over a number of « natural districts », by which it was intended to secure the representative character of the sample. An examination of the figures, however, shows that the proportion in which the various kinds of domestic animals were represented in the sample was very varied, and this indicates that in reality the sample was not so representative as was desirable. One experiment I have made, in fact, shows that a sample with more representative qualities was comparatively easy to select. In this experiment I had to make use of the information given in the printed reports of the livestock censuses of 1900 and 1890 in Austria (2). These, however, give no figures for the various municipalities, but solely for the greater administrative units, the jurisdictions. In my experiment I was thus at a greater disadvantage than the original investigators, as the use of the jurisdictions was naturally bound to give less dispersion than could be obtained by using the municipalities. In selecting the sample I proceeded as follows: — in the report of the livestock census of 1900 there is, for each jurisdiction, a calculation of the number of horses, cattle, sheep and pigs per 100 hectares of agricultural area. By successive use of these proportionate figures I divided the 219 jurisdictions into groups and sub-groups, so that finally I was able to select 48 jurisdictions where the various degrees of « animal density » (number of horses per 100 hectares, number of cattle per 100 hectares, etc.) were represented in fairly equal proportions. That by this means I arrived at both a more representative sample and a result which, in any case with regard to the most difficult point in the enquiry (the number of sheep), showed a smaller deviation from the actual figures that the original enquiry, appears from the following figures :

(1) Dr. Wilhelm Winkler: *Betriebsgrösze und Anbauverteilung. Eine kritisch-methodologische Untersuchung zur statistischen Teildarstellung.* (Jahrbücher für Nationalökonomie und Statistik. 1921. II).

(2) Oesterreichische Statistik, Vol. XXXIV, LX.

	Actual Number of animals 1900	Sample in percentage of the totality		Caiculated figures for 1890 in percentage of the actual	
		Original enquiry	My enquiry	Original enquiry	My enquiry
Horses	229,564	24.6	19.1	98.73	97.56
Cattle	2,258,338	18.1	19.8	94.10	99.12
Sheep	228,307	28.7	17.7	75.10	104.96
Pigs.	688,822	18.5	16.3	102.93	95.60

If my method of selection had been used on the municipalities, whose number is more than 30 times as large as that of the jurisdictions (by which a much wider dispersion would have been obtained), it seems very probable that the divergence between the actual and the calculated figures would fall within very narrow limits.

The method of selection used in the just described experiment and which, in principle, has also been applied to several of the representative enquiries referred to in the foregoing sections, is very well adapted indeed to create a sample which, in certain respects, possesses the same average characteristics as the totality. I think there is good reason for paying attention to the advantages offered by this method. On the other hand one must naturally not shut one's eyes to the fact that the selection of a number of units with the express purpose of forming a sample with certain average characteristics, contains a danger that the units which, in one respect or another, have extreme characteristics, will either not be represented at all or be too weakly represented. The method of selection adopted in the Danish Agricultural Census of 1923 (see section IV) has incontestably given samples which, as regards a number of important factors, show excellent conformity with the totality. But this does not exclude the possibility that the sample *can* have important defects. If, for example, it could be imagined that in some few parishes in Denmark there were no pigs at all, it is very probable that these parishes would not be represented, despite the fact that they presented peculiarities well worthy of notice.

In some cases the unfortunate result of a representative investigation has been owing to the fact that, in the selection of the sample, sufficient dispersion of the units has not been obtained.

There is a good example of this in the enquiry made by Marcus Rubin into the marriages, births and deaths in Denmark. Rubin

compiled his sample by selecting the first circuit in every county. This, however, produced a very poor dispersion and the various counties were very variously represented (from 5 to 30 per cent of their population). This was appearently the reason why the enquiry failed in regard to the number of illegitimate births. I have remade Rubin's enquiry, taking every seventh parish by mechanical counting out. Some figures are given below showing the difference between the original enquiry (A) and the one I made (B):

	Sample in percentage of the totality		Percentage of difference of calculated figures and actual figures	
	A	*B*	*A*	*B*
Population	14.98	15.33	—	—
Marriages	14.59	15.46	2.61	0.84
Total births	14.87	15.52	0.73	1.24
Illegitimate births	18.15	15.42	21.15	0.60
Deaths	15.49	15.51	3.43	1.14

An example of a representative enquiry, the less fortunate outcome of which was probably due to too little dispersion of the units in the sample, is provided in the attempt already mentioned to calculate the livestock in Prussia in 1921. Whilst Mayet, in his corresponding enquiry for Baden, used as a sample one-tenth of the country's 1,609 municipalities, Lucht worked in Prussia with 53 of the 528 « Kreise », that is to say with much greater administrative districts. How poor a dispersion was thus obtained may be seen for instance in the fact that the sample for the province of Schleswig-Holstein consisted of 2 (out of 22) « Kreises »; indeed, for the province of Grenzmark Posen–West Prussia, of only one single « Kreis » (out of 9). The greatest dispersion is naturally obtained when, as in the Norwegian representative livestock censuses, one avoids building up the sample out of administrative districts, and selects a certain fraction of the farms within each district. (On the other hand this method involves the difficulty in getting farms of various sizes represented in the right proportions).

In some cases the poor result of a representative enquiry is the outcome of the figures not being big enough to neutralise the effects of incidental causes.

In the Dutch trial-investigation into mortality, which has been referred to in Section III, no satisfactory result was expected, as it

was considered necessary to carry out the enquiry for each province separately. It also turned out that the figures were too small to give sufficiently certain results.

There is another example in Dr. Kluge's enquiry regarding the livestock in Prussia. One of the reasons why this enquiry was not successful was that Dr. Kluge calculated the results, not for the whole country or for large portions of it, but for such very limited areas as the Prussian « Kreise ».

It is possible to show in many cases that the figure material worked upon has doubtless been large enough to permit of certain general results being attained, whereas the figures have been too small when it became a question of carrying the enquiry more deeply, because when dividing up the material there often appear groups with a very small number of units. One example of this may be seen in the above mentioned English enquiry regarding the unemployed (Section II). Among the tests on the reliability of the sample which are recorded in the report, there was one which concerned the number of cases in which benefit was authorised in respect of numbers of dependent children. The test was made by comparing the figures of the one per cent sample of the November enquiry with the ten per cent sample of the January enquiry ; the figures are given below :

	Men		Women	
	November Enquiry	January Enquiry	November Enquiry	January Enquiry
Number on which benefit was authorised in respect of dependent children. . .	2,526	99,074	22	1,118
Percentage of cases in which benefit was authorised in respect of the undermentioned numbers of dependent children :				
1 child.	35.9	37.6	41.0	64.6
2 children	28.8	27.3	50.0	24.7
3 »	17.1	16.8	4.5	7.7
4 »	9.7	10.1	4.5	2.2
5 »	5.5	5.4	--	0.4
6 »	2.3	2.1	—	0.2
7 »	0.7	0.6	—	0.2
8 »	0.0	0.1	—	—
	100.0	100.0	100.0	100.0

Whilst it must be said that the test was very successful as regards the men, the report very rightly says with regard to the women that « the untrustworthiness of the sample where very small numbers are involved is apparent ».

A number of examples have been given in the foregoing showing that the unsatisfactory result of the representative enquiries has been due to some error or other committed in the planning or carrying out of the enquiry. It would not be difficult to enlarge this collection of examples by many others. I think, however, that those given are sufficient to show that one ought not be too hasty in condemning the method used in an unsuccessful enquiry. If my « Criticism of the Criticism » has also succeeded to some degree in drawing attention to some of the errors which experience shows are frequently committed in representative enquiries, so much the better.

IX. — Bibliography.

The following is an attempt to give a list of books and treatises in which the representative and similar methods are treated from a theoretical point of view.

ALTSCHUL : *Studie über die Methode der Stichprobenerhebung* (Archiv für Rassen und Gesellschafts – Biologie, 10 Jahrgang, p. 110-152. Leipzig & Berlin, 1913.

BOWLEY A. L. : *Elements of Statistics*. London, 1920.

— *An Elementary Manual of Statistics*, p. 56-64. London, 1910.

— *Address to the Economic Science and Statistics Section of the British Association for the Advancement of Science. York*, 1906. (Journal of the Royal Statistical Society, vol. LXIX). London, 1906.

BOWLEY A. L. & BURNETT-HURST A. R. : *Livelihood and Poverty*, p. 174-184. London, 1915.

BUDAY LÁSZLÓ : *A reprezentativ számlálásról*. Közigazgatasi Szemle XXII ev folyam 603-613 lap. Budapest, 1898.

— *A statisztika elmélete és története* 74-82 lap. Budapest, 1923.

BURNETT-HURST A. R., vide BOWLEY.

CZUBER E. : *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung*. I Bd., p. 203 ; II Bd., p. 12. Leipzig & Berlin, 1914.

FALKNER ROLAND P. : *The Theory and Practice of Price Statistics* (Proceedings of the American Statistical Association, 1892).

— *Wage Statistics in Theory and Practice* (Proceedings of the American Statistical Association, 1899).

FÖLDES BÉLA : *Statisztikai eloadosök* I. kötet 47, lap. Budapest, 1904.

GRAVELL W.: *Die Not der Statistik und die Repräsentativ-Methode* (Allgemeines Statistisches Archiv, 13 Band, pp. 345 ff.) – Nachwort von *Friederich Zahn*, ibid., p. 352.

— *Die repräsentative Methode*. Deutsches Statistisches Zentralblatt, 15 Jahrgang, p. 5.

HILTON JOHN: *Enquiry by Sample, an Experiment and its Results* (Journal of the Royal Statistical Society, vol. LXXXVII, p. 544).

JAHN GUNNAR: *Statistikkens Teknik og Metode*, p. 65-77. Kristiania, 1920.

JENSEN ADOLPH: *Arbejdsbesparende Metoder i Statistikken* (Nordisk Statistisk Tidsskrift, Bd. 2, p. 409-434. Stockholm, 1923).

— *Méthodes permettant de réaliser une économie de travail dans la statistique* (Présenté à l'Institut Int. de Statistique, XV Session. Bruxelles, 1923). Copenhague, 1923.

JULIN ARMAND: *Précis du cours de statistique générale et appliquée*, 5e éd., p. 25. Bruxelles et Paris, 1923.

— *Principes de statistique théorique et appliquée*. Tome I, Statistique théorique, p. 154-164. Paris & Bruxelles, 1921.

JURASCHEK: *Die IX Session des internationalen statistischen Institutes in Berlin* (Statistisches Monatsschrift, 1904).

KAUFMANN Al.: *Theorie und Methoden der Statistik*, p. 362-374. Tübingen, 1913.

KENÉZ BÉLA: *A statisztika elmélete*, 23 lap. Pozsony, 1903.

KIAER A. N.: *Den representative Undersögelsesmethode* (Skrifter udgivet af Videnskabsselskabet i Christiania, 1897. Historisk-filosofisk Klasse. No 4). Kristiania, 1897.

— *Den representative Undersögelsesmethode* (Förhandlingar vid nordiska statistiska mötet i Stockholm den 19-23 augusti 1897. Statistisk Tidsskrift, 1897, p. 93). Stockholm, 1897.

— *Die repräsentative Untersuchungsmethode*. (Allgemeines statistisches Archiv, 1899).

— *Sur les méthodes représentatives ou typologiques appliquées à la Statistique* (Bulletin de l'Institut International de Statistique, tome XI, 1e livr., p. 180).

— *Sur les méthodes représentatives ou typologiques* (Bulletin de l'Institut Int. de Statistique, tome XIII, 1e livr., p. 66).

— *Observations et expériences concernant des dénombrements représentatifs* (Bulletin de l'Institut Int. de Statistique, tome IX, 2, p. 176).

— *Statistische Beiträge zur Beleuchtung der ehelichen Fruchtbarkeit*, I, p. 126. Christiania, 1902.

LAPLACE: *Œuvres*, tome VII, no. 31, pages 430 et suivantes. Paris, 1847.

LIESSE A.: *La statistique, ses difficultés, ses procédés, ses résultats*, pp. 21-45. Paris, 1905.

LUCHT JOHANNES: *Die repräsentative Methode in der Statistik* (Zeitschrift des Preussischen statistischen Landesamts, 1922. p. 122).

MAYET P.: *Stichproben-Erhebungen in der Zwischenzeit zwischen grossen Vollzählungen längerer Periodizität* (Bulletin de l'Institut Int. de Statistique, tome XIV, 2e livr., p. 258).

v. MAYR G.: *Statistik und Gesellschaftslehre*, 1 Bd. Theoretische Statistik, 2 Aufl., p. 9-18. Tübingen, 1914.

— *Orientierung und Statistik* (Allgemeines Statistisches Archiv, 1 Jahrg., p. 429 ff.).

— *Deutsche Arbeiter-Statistik. Methodologisches und Technisches* (Allgemeines Statistisches Archiv, 3 Jahrg., pp. 131 ff.).

Meitzen August : *Geschichte, Theorie und Technik der Statistik*, Berlin, 1903.
Nybölle Hans Cl. : *Om Middelfejlen ved partielle Undersögelser* (Nordisk Statistisk Tidsskrift, Bd. 2, p. 435-36. Stockholm, 1923).
Pfaundler Richard & Weyr Franz : *Die stichprobenweisen Viehschätzungen* (Statistische Monatschrift, 1906, p. 552 & 695).
Rath Zoltàn : *Magyarország statistikája*, 40 lap. Budapest, 1896.
Rubin Marcus : *Statistiske Móder i Sommeren* 1897 (Nationalökonomisk Tidsskrift 1898, p. 101).
Sagawe : *Die statistische Methode und ihr Wert für die Wirtschaftslehre des Landbaues* ((Archiv für exakte Wirtschaftsforschung – Thünen-Archiv), 6 Bd., 1915, p. 116).
Schott Sigmund : *Beiträge zur Statistik der Stadt Mannheim*, 4 Sondernummer, p. 35-50. Mannheim 1915.
— *Das Stichprobenverfahren in der Städtestatistik* (Beiträge zur Statistik der Stadt Mannheim, N. 34, Mannheim 1917).
— *Statistik*, 3 Aufl., p. 43-45. Leipzig-Berlin, 1923).
Tschuproff A. A. : *Die repräsentative Untersuchung* (Vortrag auf der 12 Jahresversamlung russischer Naturforscher und Arzte), 1910.
Tyszka Carl von : *Statistik*. Teil I : Theorie, Methode und Geschichte der Statistik, pp. 8 & 13. Jena, 1924.
Westergaard Harald: *Statistikens Theori i Grundrids*, pp. 253-258. Köbenhavn, 1915.
— *Scope and method of statistics* (Publications of the American Statistical Association, vol. XV, 1916-17, p. 229).
— *On the Future of Statistics* (Journal of the Royal Statistical Society, vol. LXXXI, London, 1918).
Weyr F. & Pfaundler R. : vide *Pfaundler*.
Wicksell S. D. : *Contributions to the Analytical Theory of Sampling* (Arkiv för Matematik, Astronomioch Fysik, Bd. 17, No. 19. Stockholm, 1923).
Winkler Wilhelm : *Statistik*, p. 18. Leipzig, 1925.
— *Betriebsgrösse und Anbauverteilung. Eine kritisch-methodologische Untersuchung zur statistischen Teildarstellung* (Jahrbücher für Nationalökonomie und Statistik, 1921, II, p. 449).
Yule G. Udny : *An Introduction to the Theory of Statistics*, p. 250. London, 1911.
Zahn F., vide Gravell.
Zizek Franz : *Grundrisz der Statistik*, II Aufl., p. 189-193. München & Leipzig, 1924.
Comptes rendus des sessions de l'Institut International de Statistique.
Cinquième session (Berne, 1895).
Bulletin, tome IX, 2 :
Rapport : A. N. Kiaer : *Observations et expériences concernant des dénombrements représentatifs*, p. 176.
Discussion, p. xciii : *Observations par* L. Bodio, E. Cheysson, L. Guillaume, A. N. Kiaer, E. Levasseur, G. von Mayr, G. E. Milliet, H. Rauchberg, G. Schmoller.
Sixième session (St. Pétersbg., 1897).
Bulletin, tome XI, 1 :
Rapport : A. N. Kiaer : *Sur les méthodes représentatives ou typologiques appliquées à la statistique*, p. 180.

Huitième session (Budapest, 1901).

Bulletin, tome XIII, 1 :

Rapport : A. N. KIAER : *Sur les méthodes représentatives ou typologiques*, p. 66.

Discussion, p. 70 : Observations par L. VON BORTKIEWICZ, A. JULIN, A. N. KIAER. G. LANGE, P. MAYET, G. VON MAYR, G. E. MILLIET, E. MISCHLER, E. NICOLAI, Z. RÀTH, M. RUBIN, N. A. TROINITSKY, C. D. WRIGHT.

Neuvième session (Berlin, 1903).

Bulletin, tome XIV, 1 :

Discussion, p. 119 : Observations par F. VON JURASCHEK, A. N. KIAER, KLUGE, L. MARCH, P. MAYET, G. v. MAYR, K. H. ZELLER.

Bulletin, tome XIV, 2 :

Rapport : P. MAYET : *Stichproben-Erhebungen in der Zwischenzeit zwischen grossen Volkzählungen längerer Periodizität*, p. 258.