A METHOD OF ESTIMATING FUTURE POPULATIONS.

By T. H. C. STEVENSON, M.D. (Lond.), D.P.H. (Camb.).

It will perhaps be generally admitted that few of the practical problems which have from time to time to be faced by the Medical Officer of Health in his capacity of statistician present more difficulty than is often involved by the estimation of populations. This problem may be said naturally to present itself in three different forms, according as the estimate is required for a past, a present, or a future date. Practically, however, there are but two primary divisions of the problem, the one relating to dates since which a census has been taken (and its results published), and the other to dates subsequent to that of the last published census. This distinction is pointed out by Dr Cressy Wilbur in the "Thirty-second Annual Report (for 1898) on the Vital Statistics of the State of Michigan [3]." He there refers to the first class of cases as estimates of inter-censal, and the second as estimates of post-censal populations.

Inter-censal estimates are the subject of a recent paper by Dr J. Spottiswoode Cameron [1]. He compares the results for Leeds obtained by a method of interpolation with those to be got by assuming the population to have increased in geometrical progression between the years 1891 and 1901. The discrepancies are not serious, and for most purposes the simpler method of geometrical progression may be accepted as yielding sufficiently correct results.

In the case of post-censal estimations, however, the difficulties are in all cases much greater, and vary in degree with the interval between the date of the last available census and that for which the estimate is required. Three subdivisions of this class of estimates may be made, according as the estimate is required for (a) a past or present date, (b) a date in the near future, or (c) a date in the more or less remote future. A sharp line of division is drawn between groups (a) and (b)

14—2

by the fact that for group (*a*) various auxiliary methods of checking estimates of population, such as those dependent on the number of inhabited houses and on the birth-rate, can be used, which are not available for group (*b*). The methods to be used for groups (*b*) and (*c*) must also differ, but there is in this case no natural boundary-line defining the point at which the distinction must be drawn.

The method of most general application, in this country at least, to groups (*a*) and (*b*) is that of the Registrar-General, depending on the assumption of the continuance of the last ascertained *rate* of increase or decrease, and, consequently, on a future increase or decrease in geometrical progression.

Where records of migration are kept, estimates in group (*a*) are best made by adding to the natural increase the excess of immigrants over emigrants, or *vice versâ*. Thus in 1902 the population of New Zealand increased by 20,272, made up of excess of births over deaths 12,280, and excess of arrivals over departures, 7992. This method, however, can never be of general application, owing to the impossibility of tracing and recording the internal migration of any country. The method of assuming increase in arithmetical instead of geometrical progression is preferred for Michigan by Dr Wilbur in the paper referred to, and is shown to give results more closely approximating to the ascertained facts than those got by assuming increase in geometrical progression.

The Registrar-General himself, in the compilation of his published estimates, has occasion to use his method only for cases in group (*a*). But as it is used unchecked by the various other methods available for this group of estimates, the distinction in the case of his returns between groups (*a*) and (*b*) may be said to disappear. The method of estimating a population for the year 1908 which would be used in 1904, the case then coming under group (*b*), is the same as that which will be used for the returns in 1908, when the case will come under group (*a*). The circumstances of the case render it unnecessary for this method to be used in these returns for any date more than ten years subsequent to that of the last preceding census, consequently the example of the Registrar-General cannot be appealed to as authorising its use for any more distant date.

Even with this limitation the method at times yields very misleading results. Thus at the census of 1891 the population of Salford was found to be 20·9 per cent. below the estimate based on the rate of increase in 1871—81 [2]. The period within which the method may be used with some degree of confidence will vary with the circumstances of

the particular locality, but the facts above referred to seem to suggest that there is no ground for assuming the applicability of the method for any date more than ten years subsequent to that of the last preceding census. Its use for more distant dates can easily be shown to be in many cases quite out of the question. If, as is not unfrequently the case, an estimate is required of what the population of a suburban district adjacent to a large town is likely to amount to in 20 or 30 years, this method will frequently furnish results which are absurdly high. When such districts are first opened out for building purposes the intercensal rate of increase may be very high, as much as 400 per cent. in some instances. In such cases the continuance of an equal rate for even 20 years longer might result in a population denser than that of the central part of the city, to the overflow of which the filling up of the outlying district is due. Obviously a method cannot be relied upon which is apt, in a comparatively short time, to make the suburbs more densely populated than the centre of the city, or perhaps than the maximum to which density of population ever attains in this country.

It must be recognised then that, while the accepted methods give more or less satisfactory results for estimates relating to the past, the present, or the immediate future, the only one of them which from its nature is applicable to the future at all cannot be relied upon for an estimate relating to a period more than a few years ahead. If in view of this we seek for further guidance than the last ascertained rate of increase, we find that as it was the density of population obtained by this method which demonstrated its inapplicability, so in devising a better method the influence of increasing density ought if possible to be taken into account.

Doubtless other circumstances than the density of population will modify the rate of increase, such as trade conditions, transit facilities, etc., influencing migration mainly, and alterations in the birth and death rates, influencing the natural rate of increase. Although we may probably look forward to a continuance, for some time longer at least, of the present steady fall in birth and death rates, still as the changes occurring tend to neutralise each other in their effect upon natural increase of population, it seems best not to attempt any correction for anticipated alterations in these rates. The conditions influencing migration twenty years hence, apart from density of population, cannot be anticipated at the present time, so alterations in these factors cannot be allowed for; in other words we are bound to assume the continuance of past conditions because we cannot in any degree foresee their future

alterations. The one condition of which the change can to a certain extent be anticipated and allowed for is density of population. Consequently the method to be described assumes, *faut de mieux*, the continuance in all other respects of past conditions, but modifies the results obtained in order to allow for the influence of increasing density of population.

In view of the very misleading results obtainable by the application of the method of geometrical progression in such cases Dr Thresh, Medical Officer of Health of the county of Essex, who required an estimate of the probable population in 1921 and 1931 of certain districts in this county bordering upon the county of London and the county borough of West Ham, suggested that in framing it advantage might be taken of the past experience of West Ham and East London. The present paper describes a method devised for the purpose of utilising this suggestion.

The districts of greater London to be estimated for were East Ham, Barking, Ilford, Leyton, Wanstead, and Walthamstow.
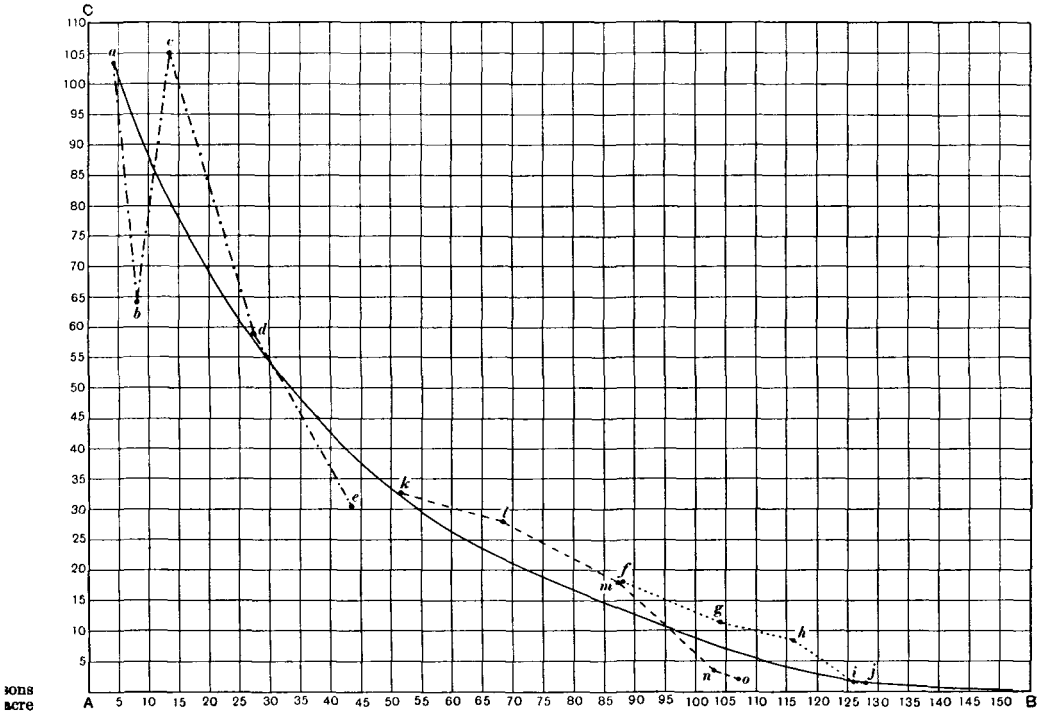
They have been filling up at an exceedingly rapid rate in the recent past, and in the case of some it would be absurd to anticipate the continuance till 1931 of an equal rate of increase. Taking the case of East Ham for instance, the population here was 32,713 in 1891 and 96,018 in 1901. The continuance for the next thirty years of an equal rate of increase would produce a population in 1931 of about 2,428,000, or 742 persons per acre, whereas the density of population in Whitechapel in 1901 was about 207 persons per acre.

As the districts in question are contiguous, and, on the whole, similar in character, it has been thought best to treat them as a whole instead of estimating separately for each, and they will therefore be referred to here for the sake of convenience, as the "combined districts."

The neighbouring borough of West Ham was selected to furnish guidance as to the probable course of their further development, as it may be said to have occupied thirty years ago the developmental position in which the "combined districts" find themselves to-day, and as the type of district is on the whole similar. As, however, the West Ham figures do not indicate the probable rate of increase beyond the point at which a density of 43·5 persons per acre has been attained, they have been supplemented by those for the "East End" Registration Districts of the county of London, of which, taken as a whole, the density in 1851 was 88, and in 1901 130 persons per acre. The

CURVE OF RELATIONSHIP OF INCREASE OF POPULATION TO DENSITY IN EAST LONDON 1851—1901.

	$a=$	percentage increase (103·7) in West Ham population during	1851—61, the density in 1851 being 4 persons per acre.								
	$b=$	,,	,,	(64·1)	,,	,,	,,	1861—71	,,	1861 ,,	8·1 ,, ,,
	$c=$	,,	,,	(105·0)	,,	,,	,,	1871—81	,,	1871 ,,	13·4 ,, ,,
	$d=$	,,	,,	(58·9)	,,	,,	,,	1881—91	,,	1881 ,,	27·4 ,, ,,
Figures, upon which curve is founded	$e=$	,,	,,	(30·5)	,,	,,	,,	1891—1901	,,	1891 ,,	43·5 ,, ,,
	$f=$	,,	,,	(18·2) in East End Registration Districts	1851—61		,,	1851 ,,	88 ,,	,,
	$g=$	,,	,,	(11·5)	,,	,,	1861—71	,,	1861 ,,	104 ,, ,,
	$h=$	,,	,,	(8·6)	,,	,,	1871—81	,,	1871 ,,	116 ,, ,,
	$i=$	,,	,,	(1·6)	,,	,,	1881—91	,,	1881 ,,	126 ,, ,,
	$j=$	,,	,,	(1·4)	,,	,,	1891—1901	. ,,	1891 ,,	128 ,, ,,
Figures, not used for construction of the curve, but applied afterwards as a test	$k=$	,,	,,	(32·7) in {Bethnal Green, Mile End and Poplar}	,,	1851—61	,,	1851 ,,	51·5 ,, ,,
	$l=$	,,	,,	(28·1)	,,	,,	1861—71	,,	1861 ,,	68·4 ,, ,,
	$m=$	,,	,,	(18·0)	,,	,,	1871—81	,,	1871 ,,	87·6 ,, ,,
	$n=$	,,	,,	(3·7)	,,	,,	1881—91	,,	1881 ,,	103·3 ,, ,,
	$o=$	,,	,,	(2·0)	,,	,,	1891—1901	,,	1891 ,,	107·2 ,, ,,

numbers of persons per acre in West Ham and in the East End Registration Districts at each census from 1851 onwards have been ascertained, as well as the corresponding percentage increase of population occurring in the ensuing decennium. Thus in 1881 the density of population in West Ham was 28·6 per acre, and the rate of increase in the following ten years was 58·9. The method adopted assumes that when the density of the "combined districts" amounts to 28·6, their rate of increase in the following ten years will also approximate to 58·9 per cent. Having obtained the rates of increase corresponding to ten individual density figures (five censuses, 1851—1891, and two communities being used) it was necessary to deduce from them the probable rate of increase corresponding to any given density figure.

The graphic method of interpolation was used for this purpose, and the curve of rates of increase obtained is represented in the figure, p. 211. The method of its construction is as follows: The different degrees of density of population, measured in persons per acre, from 0 to 150, are marked off along the abscissa line AB, and on the line AC, parallel to the ordinates, a scale of percentage decennial increases is marked. Then the different percentage increases of population recorded for West Ham and for the East End Registration Districts are charted vertically above the figure on the abscissa line AB representing the density of population recorded at the census marking the commencement of the ten yearly period during which the increase occurred. For instance, the West Ham increase of 58·9 per cent. during 1881—1891 is charted at the point marked d above the point on the line AB corresponding with the density of population of West Ham in 1881, namely 28·6. When this is done in each case it is found that, with one exception in the case of West Ham, the percentage rate of increase falls as the density increases. From the two sets of points thus determined, which are united by continuous black lines in the diagram, a curve has been drawn, filling in the gap between them (between densities 43·5 and 88) and smoothing off the irregularities met with, which, except in the instance referred to, relating to the early development of West Ham, are not very great. From this curve a rate of increase for the following ten years can be read off corresponding to any given density, and by means of it the estimate for the "combined districts" was prepared. Thus their density at the 1901 census was 14·9 persons per acre, and the rate of increase determined by the curve to correspond to this density is 78 %. At this rate the 362,021 inhabitants of the "combined districts" would

become 644,397 in 1911. This is equivalent to a density of 26·6 per acre, and the corresponding increase as determined by the curve is 59%, which gives a population in 1921 of 1,024,591.

The simple curve obtained by following merely the general trend of the recorded figures is used rather than the irregular curve which would result from following the recorded experience in every case, because it seems likely that by doing so the influence of circumstances other than density of population is more or less eliminated. If the density were the only influence at work in determining the rate of increase we must suppose that it would act steadily, not irregularly, and that some such simple curve as is here constructed would represent its effect. If that is so then the irregularities of the series of recorded figures must be due to variations from time to time in local conditions other than density of population, and therefore are, for our purposes, best disregarded. But while it seems reasonable to suppose that a curve expressive solely of the influence of density would be simple, it by no means follows that, if it could be accurately ascertained, it would correspond with the curve in the figure. All that is claimed for the latter is that it represents an attempt to express the general trend of the recorded figures, and that no curve differing very widely from it would do so, or would be likely to represent the influence of density.

The estimates obtained from the curve, 1,024,591 for 1921 and 1,454,919 for 1931, may seem very high for districts containing in 1901 only 362,021 souls. London has however in recent years over-flowed with great rapidity into these districts. Figures for East Ham have already been quoted. Its population was 10,706 in 1881, 96,018 in 1901, and is now estimated at about 115,000. Some of the other districts give figures almost as remarkable. If the method of a continuance in geometrical progression of the recent rate of increase is applied to the "combined districts" as a whole the estimate for 1921 becomes 1,558,869, an excess of over half a million over that obtained by the method adopted. This difference illustrates the extent to which an estimate based only on previous rate of increase is modified when the influence of increasing density is taken into account.

On the other hand, in the absence of a rational method of estimation, such as that founded on the basis of density is believed to be, the probable increase of population is liable also to be greatly under-estimated. For instance, West Ham and other portions of the districts now dealt with have been said recently to be "full up," and not likely to increase in population. The truth of such a statement is best tested by

comparison with the experience of other similar districts, that is, on the density basis. When this is done we find that the density of West Ham in 1901 was 56·8, and that the three adjoining districts of Bethnal Green, Mile End, and Poplar had a combined density in 1861 of 68·4, with a percentage increase in the following decade of 28·1. Nor can it be said that conditions have altered so much for the better since that time that the analogy is inapplicable. The density of the East End Registration Districts was 126 in 1881, but an increase to 128 in 1891 and to 130 in 1901 occurred notwithstanding. Recent rehousing estimates for the London County Council have assumed a density of 200 per acre.

The above three have been selected from amongst the seven East End Registration Districts in order to test the curve already obtained as above described. Of the seven districts these three alone have grown to any extent since 1851, and so we may take it that the other four, which had a density of 168 persons per acre in 1851, and 175 in 1901, had almost reached the limit of density at the earlier date. The three selected, however, lying further from the City than the other four, had not reached nearly the same stage of development in 1851. Their density at that time was 51·5, and in 1901, 109·3. The East End Registration Districts therefore may be divided into an inner and an outer group, the history of the outer, as represented by the dotted lines in the figure, furnishing a test, to a certain extent, of the accuracy with which the graphic method has enabled the gap left between the records of West Ham, which cease at a density of 43·5, and those of the East End Registration Districts as a whole, which commence at 88, to be filled in. The increase in these three districts, following their density of 51·5 in 1851, namely 32·7, almost exactly corresponds with that read off the curve. The 1861—71 and 1871—81 increase rates are a little above, and the 1881—91 and 1891—1901 rates about the same amount below, those given by the curve for the corresponding densities. It follows that if these figures had been ascertained before the curve was drawn its form would not have been altered, as it continues to express the average of all the data secured.

The particular case dealt with is no doubt an especially favourable one for the application of the method described, for it has been possible to secure instances of neighbouring and approximately similar districts at all stages of development, but it is believed that the materials for its application exist in many other cases, though generally, perhaps, in a less complete form. When it is desired in any case to estimate the

probable increase of population in any outlying suburb there must often be some inner suburb whose past record can be used for the construction of such a curve as that described. When this can be done it is anticipated with confidence that the result will be more reliable than any to be obtained by the method of guess-work, however discreet and well-informed, or by assuming the maintenance of a past rate of increase, provided that the condition of similarity between the populations is borne in mind. It will be understood of course that only the third group (c) of post-censal estimates is being discussed, and that the method is not recommended for estimating populations of a few years hence. Nor does it seem likely to be useful for estimating the future populations of rural districts or of the central parts of large towns. These, however, are generally much more stationary than the suburban districts, and it is in the case of the latter, where the estimates are of greatest importance and greatest difficulty, that it is believed that substantial assistance will be afforded by the method described.

It will now be apparent why the method of arithmetical progression may well give better results, as tested by recorded facts, than that of geometrical. The latter method entirely ignores the influence of increasing density, and the curve, analogous to that in the figure, that must be traced if it be relied on, is represented by a horizontal straight line (the rate of increase remaining the same whatever be the density). The method of arithmetical progression on the other hand involves the assumption of a rate of increase decreasing as the population becomes more dense, and therefore, for communities whose space is limited, will more nearly approximate to the truth in the long run. The limitation of space is a necessary condition however. If we wish to estimate the future population of a borough which will from time to time enlarge its boundaries we shall probably best ascertain it, for the borough as enlarged, by geometrical progression. If, on the other hand, we wish to ascertain the future population of a definite restricted area, the use either of some such method as here described, or of the method of arithmetical progression, should generally give better results than geometrical progression. Similarly, if a country or a province is mainly agricultural the element of density comes in, since the amount of land available for crops or pasture is strictly limited, and, given free migration, the method of arithmetical progression would probably be preferable. But if the country or province is industrial, the space available being practically unlimited, the element of density does not

come in to any great extent, and as in the case of England, the method of geometrical progression may be expected to yield good results.

In conclusion, reference may be made to two obvious disadvantages of the "density method" described. The experience relied on is not that of the district itself but of some other, though similar, district, and the experience used may be many years old. These are both grave disadvantages, and a method involving them must be, at best, unsatisfactory. The nature of the problem, however, is such that a satisfactory solution is impossible, and the method described has seemed, in the particular instance given, to be less unsatisfactory than any other which suggested itself.

REFERENCES.

(1) CAMERON, J. S. (x. 1903).   The estimation of intercensal populations.   *Public Health*, vol. XVI. p. 34.
(2) NEWSHOLME, A. (1899).   *Vital Statistics*, 3rd edition (London : Swan Sonnenschein and Co.), p. 8.
(3) NEWSHOLME, A. (1901).   Methods of estimating population for the purposes of vital statistics.   *Practitioner*, vol. XIII. n. s. p. 428.