AN EXAMINATION OF SOME FACTORS INFLUENCING THE RATE OF INFANT MORTALITY.

By M. GREENWOOD, Jr.,
Statistician to the Lister Institute of Preventive Medicine.

and J. W. BROWN,
Assistant in the Statistical Laboratory, The Lister Institute of Preventive Medicine.

I. *Introduction.*

In all countries which make statistical returns, attention has been directed to the large proportion of children failing to complete one year of life. Even the most highly civilised races lose in this way more than ten per cent. of their potential citizens as will appear from the data contained in the following table (*Deutsche Sterbetafeln*, 1910).

Country	Number out of 100,000 born who survive the first year
Sweden	88,917
France	83,674
Belgium	83,114
England and Wales	82,814
Massachusetts	82,767
Holland	82,681
Italy	82,481
German Empire	76,614
Austria	75,028

The significance of this phenomenon, its advantage or disadvantage to the state, has been variously interpreted by writers on the subject. Some have argued that a high rate of infant mortality may be of advantage, that persons ill-adapted to their social and physical environment are in this way eliminated with a minimum of individual suffering and material inconvenience to the community. Others, we think the

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

majority, have taught that a high infant death rate is detrimental, that the elimination even if selective does not necessarily remove those calculated to prove inferior citizens.

Whether the infant death rate is selective and, if selective, whether such a process is really beneficial to the state are questions we do not propose to discuss at length, although some slight reference to the subject will be unavoidable in connection with the concluding passages in our memoir; we shall be mainly occupied with certain preliminary topics. If the factors determining the rate of infant mortality are within our control, the consequences of exercising that control may well be considered; if, on the other hand, the factors are beyond the scope of human interference, such a discussion would be of purely speculative interest. At one time, it was thought that a high infant death rate was part of the established order of nature, as it were, inherent in the necessary constitution of society. In the famous essay of Malthus (1817) this view finds adequate expression. Thus, he writes (Vol. II. p. 49) "And yet when we contemplate the insufficiency of the price of labour to maintain a large family, and the amount of mortality which arises directly and indirectly from poverty; and add to this the crowds of children, which are cut off prematurely in our great towns, our manufactories and our workhouses; we shall be compelled to acknowledge that if the number born annually were not greatly thinned by this premature mortality, the funds for the maintenance of labour must increase with much greater rapidity than they have ever done hitherto in this country in order to find work and food for the additional numbers that would then grow up to manhood."

According to Malthus fifty per cent. of all persons born died in London under the age of three years, in Vienna and Stockholm under two, in Manchester and Norwich under five and in Northampton under ten (Vol. II. 58). In the whole of Sweden the average mortality in the first year was 25 per cent. (I. 421). (For a discussion of older natality and mortality statistics, see Wernicke, 1889.)

At the present time, heavy as is the toll taken of infant lives, it is markedly less than indicated in the figures just mentioned and most writers have attributed the reduction to improved environmental conditions, using that word in a broad sense. We must, however, remember that not only has sanitary science advanced but, in addition, the birth rate has fallen. Were Malthus alive now, he would probably attribute a large share in the diminution of the Infant Death Rate to the fall in the Birth Rate and it is important to learn how far such an attribution would be correct, since the Birth Rate is hardly amenable to direct administrative control.

Numerous collections of data have been published with the object of determining which of the various factors may be responsible for the observed result; it is to this question which we desire to devote ourselves.

Since what may be called the environmental factor is capable of resolution into many components, it will be seen that the problem is too complex for any satisfactory solution to be reached without a minute and somewhat difficult analysis. It has seemed to us probable that an application of modern statistical methods to existing collections of data might lead to results suggesting further lines of inquiry, and the present essay is the result of such an attempt. Evidently no single group of workers can possibly deal with a tithe of the accumulated facts, but any contribution, however slight, may be of service to those called upon to face one of the most serious problems falling within the scope of Preventive Medicine.

II. *Nature of the Material Studied.*

Our work is based upon a statistical reduction of certain figures contained in an important memoir by Dr Alfred Groth and Prof. Martin Hahn, entitled *Die Säuglingsverhältnisse in Bayern*[1]. We shall therefore first describe the nature of these authors' inquiry, the material they collected and the conclusions they were led to adopt. The authors set themselves the task of elucidating the connections between Infant Mortality, the Methods of Infant feeding, the Birth Rate, the Degree of Poverty and the Fitness of Recruits for Military Service in different parts of the kingdom of Bavaria. In their first section they publish tables giving the population, the birth rate and the death rate of infants for the provinces of Bavaria from 1835–1904. From these figures various conclusions are drawn which do not concern us for the moment. They then pass on to the state of affairs in Bavaria between 1900–4. This part of the memoir deals with the inter-relations of Infant Mortality, Artificial Feeding of Infants, the Birth Rate and the Poverty Rate. For the purposes of examination the data are grouped under Regierungsbezirke (Oberbayern, Niederbayern, Pfalz, Oberpfalz, Oberfranken, Mittelfranken, Unterfranken, Schwaben) and a further

[1] Cited in future as Groth and Hahn.

subdivision is effected into Unmittelbaren Städte and Bezirksämter[1]. With regard to this latter distinction, the authors write (p. 14) "Die unmittelbaren Städte, die zwar von verschiedener Grösse doch nur ausnahmsweise weniger als 5,000, in den meisten Fällen wesentlich mehr Einwohner besitzen, treten einmal durch das ihnen eigene engere Zusammenwohnen auf kleinerem Terrain und dann durch die Zusammensetzung ihrer Bewölkerung aus verschiedenen sozialen Schichten in Gegensatz zu den das Land in engeren Sinne darstellenden Bezirksämtern. Wenn auch die letzteren sehr häufig Städte mit einigen tausend Bewohnern in sich beschliessen, so lassen dieselben doch mehr den Charakter des sie umgebenden flachen Landes durch ihre mehr oder weniger weitgehende Abhängigkeit von demselben erkennen."

The exact degree of differentiation here indicated cannot be profitably discussed by foreign critics. Thus, it is not clear how the delimitation is effected, *e.g.* we have Unmittelbare Stadt München with a population of 519,162 and the Bezirksamt München with 44,092. This hardly suggests that the Bezirksamt corresponds to a suburban zone unless the tendency to the decentralisation of dwellings on the part of urban populations, which is so marked in England, does not exist in Bavaria. On the whole, however, it may reasonably be presumed that a Bezirksamt corresponds to an English Rural District and an Unmittelbare Stadt to what we term an Urban District.

We now come to the variables investigated.

(1) *Population.* "Die mittleren Bevölkerungszahlen der einzelnen Verwaltungsbezirke wurden dadurch gewonnen, dass einmal die für das Volkszählungsjahr 1900 erhobenen und in dem Gemeindeverzeichnisse des Königreichs Bayern niedergelegten Zahlen und dann die des darauffolgenden Zählungsjahres, nähmlich des Kalenderjahres 1905 zur ihrer Berechnung herangezogen wurden" (p. 14).

(2) *Birth Rate.* This is the number of live-born children per cent. of the whole population (1900–1904).

(3) *Infant Death Rate.* The number of deaths in the first year of life per hundred live-born children (1900–1904).

(4) *Poverty.* Consultation with the officials of the Finance Ministry having convinced the authors that no satisfactory measure could be derived from taxation statistics, they decided to use the percentage of persons in receipt of public assistance (aus öffentlichen Mitteln unterstützte Personen).

[1] In the Pfalz, the sixteen administrative districts are not separated into towns and country districts neither are poverty statistics (*vid. infra*) given.

Groth and Hahn frankly admit that this method of estimating the economic state of a district is open to serious objections. Thus, it might well be that the extent of public assistance provided depends to some extent on the amount of money which can be raised locally, so that a rich district may give more assistance than a poor one although the latter has proportionally a greater number of necessitous inhabitants. In mitigation, they urge that this factor will rather influence the *amount* expended than the *number of persons* relieved (Groth and Hahn, p. 25). Another objection is that the poverty statistics are based on the years 1899–1902, not on the period (1900–4) to which the birth and death rates refer. That the criterion of poverty is insufficient will be recognised, but, as will be shown later on, constants based upon it bear comparison with results obtained in a better way.

(5) *The method of feeding.* Records of the extent to which the use of artificial food prevailed were obtained in a simple way. Inquiries were directed to the officers in charge of the public vaccination departments and a census made of the form of nourishment provided for the infants brought up to be vaccinated. In this way information was obtained respecting a large sample of infants in each district. Thus the smallest total upon which a percentage figure is based (in Bezirksämter) was 160, the largest 2048, and the average 628.

No doubt this method is not calculated to yield complete information regarding the prevalence of artificial feeding since, *inter alia*, it may be presumed that a considerable proportion of the wealthier inhabitants do not bring their children to the public vaccinator at all. At the same time, the value of such extensive sampling from the standpoint of the statistician and publicist is considerable. Unfortunately, as in the case of the poverty statistics, the years of record do not correspond to those used for the birth-rate, nor are they the same in all districts (1905–8). A more formidable objection, deduced from statistical considerations, will be mentioned below.

The methods adopted by Groth and Hahn in reducing their material would not altogether commend themselves to an English statistician. Much use is made of summarising tables, of which the following is an example (Groth and Hahn, p. 27).

Diagrams have also been constructed to show the relations between the variables, but are not of a nature permitting accurate comparison.

In the concluding section of their paper, Groth and Hahn compare (1) Percentage mortality in the 1st year of life; (2) Percentage of Infants not breast fed; (3) Percentage of recruits examined and declared

Poverty and Infant Mortality, mean value per 29 *Bezirksämter arranged according to the number of paupers,* 1899–1902.

	Teuschnitz to Aschaffenburg	Mellrichtstadt to Münchberg	Eggenfelden to Günzburg	Hersbruch to Hilpoltstein	Dillingen to Graffenau
Mean Poverty-Figure	1·38	1·88	2·32	2·79	3·65
Mean Infant Mortality-Figure	20·9	20·8	26·4	27·4	30·4

fit; (4) Mortality at ages 6–10; (5) Mortality at ages 11–20. Comparisons are effected as in the earlier chapters of the work.

We shall now briefly summarise the main conclusions which Groth and Hahn adopt on the basis of their enquiry.

The height of the infant mortality rate depends on the combined working of many factors, the separate influence of which ought neither to be under- nor over-emphasized. In particular, it is inaccurate to delimit a small group of so-called primary factors and to contrast them with secondary influences. "The complete examination of the infantile mortality of a race cannot be separated from the coincident study of the height of the birth rate and the number of children which survive the second year through the inter-play of birth and death" (Groth and Hahn, p. 81).

"The power of procreation does not determine the height of the birth-rate but, in the first place, the effective will of the people, increasing with improved material conditions and social insight, to limit the number of children (artificial limitation)" (*op. cit.* p. 81). The authors regard prolonged lactation as an anti-conceptual influence of power commensurate with the "heutzutage weitverbreitete Herabsetzung der Kinderzahl auf unnatürlichem Wege." They consider that the fall in the birth rate of recent years in Bavaria is to be attributed to the improvement of hygienic and social conditions which have accompanied it.

The writers further hold that breast feeding can almost neutralise (so gut wie vollkommen zu paralysieren) the influence of an unfavourable economic environment on the rate of infant mortality. They emphasize the necessity of caution in estimating the nature of the connection between infant mortality and its supposed factors, remarking that the relation between the infant birth and death rates is "too complex and too indirect to allow of its being set out in a general stereotyped way." Groth and Hahn would appear to regard the method of feeding as the most influential *single* factor at work, in particular as of more importance than general hygienic measures. They also find a low infant death rate and a high percentage of breast fed children associated with a high degree of military fitness, although the correspondence was not invariable.

It must, we think, be recognised that Drs Groth and Hahn have made an important contribution to the literature of their subject. It is, indeed, because their material, whatever may be said of its shortcomings, seems peculiarly far-reaching that we have made it the foundation of our analysis. Where, in our judgment, the paper needs extension is in the analysis of details; since what the hygienist requires is exact information regarding the relative importance of the factors displayed. Neither by summarising tables nor by diagrams is it possible to extract from the figures all the information which they should be compelled to furnish; a more refined statistical calculus is imperatively required[1].

We do not suggest that the method about to be described is perfect or that the conclusions we have reached supersede those of Groth and Hahn. We believe, however, it will be found that we have been able to place some important results in a clearer light than was possible with the simpler methods employed by the original authors.

III. *Analysis of the Bavarian data.*

We now pass to a more detailed analysis of Groth and Hahn's material. Confining ourselves in the first instance to the subject of infant mortality without referring to its effect upon adult fitness, which cannot indeed be deduced from the figures, we have four variables to consider: (1) The Infant Mortality Rate, (2) The Birth Rate, (3) The Poverty Rate, (4) The Artificial Feeding Rate. We have already indicated some criticisms to which the actual data are subject. It is at once clear that a distinction must be drawn between rural and urban districts, that the data must be divided into two classes. Our first proceeding was to calculate the correlations between the variables, taken in pairs, for the two classes separately.

Table I contains the preliminary results obtained[2].

[1] On a subsequent page we call attention to uncertainties regarding the material suggested by our examination.

[2] In working, more decimal places were retained (not less than five) than might perhaps be inferred from the tabled results.

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

TABLE I.

Crude Correlations (Bavaria).

TOWNS.

Variables		No. of observations	Coefficient of correlation
Birth Rate and Artificial Feeding Rate	...	20	– ·258 ± ·141
Poverty and Artificial Feeding Rates	...	20	·324 ± ·135
Birth and Poverty Rates	...	42	·045 ± ·104
Birth and Infant Mortality Rates ...	...	42	– ·004 ± ·104
Poverty and Infant Mortality Rates	...	42	·122 ± ·103
Artificial Feeding and Infant Mortality Rates		20	·770 ± ·062
RURAL DISTRICTS.			
Birth and Artificial Feeding Rates	...	97	·397 ± ·058
Poverty and Artificial Feeding Rates	...	87	·369 ± ·062
Birth and Poverty Rates	...	145	·423 ± ·046
Birth and Infant Mortality Rates ...	...	161	·719 ± ·018
Poverty and Infant Mortality Rates	...	145	·465 ± ·042
Artificial Feeding and Infant Mortality Rates		97	·761 ± ·029

The probable errors of the coefficients being, in the case of the urban districts, very large, little importance can be attached to the results: it would appear, however, to judge for example from the Birth Rate-Artificial Feeding coefficients, that the groups are not comparable. In what follows, attention will be exclusively directed to the Bezirksämter. (There are numerous papers especially devoted to urban Infant Mortality, *e.g.* Silbergleit, 1897.)

TABLE II.

Variables			Coefficient of correlation
Population and B. R.	...	...	·146 ± ·053
Population and Pov. R.	...	...	·091 ± ·056
Population and Inf. M. R.	...	...	– ·068 ± ·053
Population and Art. F. R.	...	...	– ·080 ± ·068

TABLE III.

Total Coefficients for the Rural Districts corrected for Population.

Variables			Coefficient of correlation
B. R. and Art. F. R.	...	...	·414 ± ·057
Pov. R. and Art. F. R.	...	...	·379 ± ·062
B. R. and Pov. R. ...	...	...	·416 ± ·046
B. R. and Inf. M. R.	...	...	·738 ± ·017
Pov. R. and Inf. M. R.	...	...	·475 ± ·042
Art. F. R. and Inf. M. R.	...	...	·760 ± ·029

To begin with, it has probably occurred to the reader that the question of spurious correlation arises, owing to the fact (*inter alia*) that the absolute population differs extremely in the various districts. It seemed, therefore, desirable to eliminate so far as possible this cause of error by calculating the correlation between absolute size and each variable, then determining the (partial) correlation with the population factor constant. Table II contains the values of the correlation between absolute size and each rate. Table III shows the result for each coefficient of keeping the population constant.

The next step was to determine the *net* or partial correlation between the variables taken in pairs for a constant value of the other two characters. In working out these coefficients, the notation introduced by Yule (Yule, 1907 and 1911) is very satisfactory and permits the arithmetical work to be so arranged that a rapid check upon its accuracy is secured. We have accordingly in every case involving four variables used this method. The results so far attained appear in Table IV.

TABLE IV.

Partial Correlations: Bavarian Rural Districts (Population Constant).

Variables	Constants	Correlation
Birth Rate and Art. Feeding Rate	Poverty Rate and Infant Mort. Rate	$-\cdot 341 \pm \cdot 064$
Poverty Rate and Art. Feeding Rate	Birth Rate and Infant Mort. Rate	$\cdot 074 \pm \cdot 072$
Birth Rate and Poverty Rate ...	Art. Feeding Rate and Infant. Mort. Rate	$\cdot 128 \pm \cdot 071$
Birth Rate and Infant Mort. Rate	Poverty Rate and Art. Feeding Rate	$\cdot 687 \pm \cdot 038$
Poverty Rate and Infant Mort. Rate	Birth Rate and Art. Feeding Rate	$\cdot 136 \pm \cdot 071$
Art. Feeding Rate and Infant Mort. Rate	Birth Rate and Poverty Rate ...	$\cdot 721 \pm \cdot 035$

The next step in our investigation was to determine whether the coefficients could be substantially altered by excluding the returns from the Bavarian Palatinate (Pfalz) since in the case of this province there is no separation into rural and urban districts. The whole of the previous calculations were accordingly repeated on the data reduced in this way (Table V). It will be noticed that some of the absolute values have been changed but not to an extent seriously affecting the interpretation of Table IV.

We next faced the difficulty introduced by the fact that Groth and Hahn calculated the birth rate upon the total population, not upon the numbers of women at fertile ages in the districts. Since it might be that marked local differences in age or sex constitution existed, this proceeding was open to criticism. It was not possible, with the aid of any published material, to remedy the defect, but—by the courtesy of Dr Zahn, Director of the Royal Bavarian Statistical Office, to whom we desire to tender our hearty thanks—we have been supplied with the numbers of women aged 16–50 enumerated in each district for the

TABLE V.

Correlation between Infant Mortality Rate, Poverty Rate, Birth Rate and Artificial Feeding Rate in Bavarian Rural Districts excluding Bavarian Palatinate (Pfalz).

Variables	No. of obs.	Correlation
Birth Rate and Art. Feeding Rate ...	87	·428 ± ·059
Birth Rate and Poverty Rate ...	145	·423 ± ·046
Poverty Rate and Art. Feeding Rate	87	·369 ± ·063
Birth Rate and Infant Mort. Rate ...	145	·790 ± ·014
Poverty Rate and Infant Mort. Rate	145	·465 ± ·042
Art. Feeding Rate & Infant Mort. Rate	87	·727 ± ·034

Correlation of Population with each of the Variables.

Birth Rate and Population	·157 ± ·056
Poverty Rate and Population	·091 ± ·056
Infant Mortality Rate and Population	·250 ± ·052
Artificial Feeding Rate and Population	·204 ± ·069

Total Correlations with Population Constant.

Birth Rate and Art. Feeding Rate	·410 ± ·059
Birth Rate and Poverty Rate	·415 ± ·046
Poverty Rate and Art. Feeding Rate	·359 ± ·063
Birth Rate and Infant Mortality Rate	·785 ± ·012
Poverty Rate and Infant Mortality Rate ...	·459 ± ·042
Art. Feeding Rate and Infant Mortality Rate ...	·713 ± ·036

Partial Correlations with Population Constant.

Variables	Constants	Correlation
Birth Rate and Art. Feeding Rate	Poverty Rate and Infant Mort. Rate	− ·354 ± ·063
Birth Rate and Poverty Rate ...	Art. Feeding Rate and Infant Mort. Rate	·126 ± ·071
Poverty Rate and Art. Feeding Rate	Birth Rate and Infant Mort. Rate	·092 ± ·072
Birth Rate and Infant Mort. Rate	Poverty Rate and Art. Feeding Rate	·746 ± ·032
Poverty Rate and Infant Mort. Rate	Birth Rate and Art. Feeding Rate	·111 ± ·071
Art. Feeding Rate and Infant Mort. Rate	Birth Rate and Poverty Rate ...	·676 ± ·039

census year 1900[1]. Since the births were averaged from 1900 to 1904, a rate calculated on the number of women recorded in 1900 would not be accurate. However, unless rapid and heterogeneous changes have taken place in the districts, such a rate should be sufficient for the present object. In Tables VI and VII the results of re-calculation upon this basis are recorded. On examination of the partial correlations it will be seen that the *relative* values are little affected.

TABLE VI.

Correlation of Infant Mortality, Artificial Feeding, Poverty and Birth Rates in Rural Districts of Bavaria including Bavarian Palatinate (Pfalz).

Birth Rate based on number of women aged 16–50.

Variables	No. of observations	Correlation
Birth Rate and Artificial Feeding Rate ...	92	·462 ± ·055
Birth Rate and Poverty Rate	142	·426 ± ·046
Birth Rate and Infant Mortality Rate ...	156	·696 ± ·028
Poverty Rate and Artificial Feeding Rate ...	87	·369 ± ·062
Poverty Rate and Infant Mortality Rate ...	145	·465 ± ·044
Artificial Feeding Rate and Infant Mortality Rate	97	·761 ± ·029
Correlation between Population and each of the Variables.		
Birth Rate and Population	156	·105 ± ·053
Poverty Rate and Population	145	·091 ± ·056
Infant Mortality Rate and Population ...	161	− ·068 ± ·053
Artificial Feeding Rate and Population ...	97	− ·080 ± ·068
Total Correlations: Population Constant.		
Birth Rate and Artificial Feeding Rate ...	92	·474 ± ·055
Birth Rate and Poverty Rate	142	·420 ± ·047
Birth Rate and Infant Mortality Rate ...	156	·709 ± ·027
Poverty Rate and Artificial Feeding Rate ...	87	·379 ± ·062
Poverty Rate and Infant Mortality Rate ...	145	·475 ± ·043
Artificial Feeding Rate and Infant Mortality Rate	97	·760 ± ·029

For the purposes of Tables V–VII, the partial coefficients have been based upon the total correlations deduced from all the available data, the probable errors being calculated on the assumption that $n=87$ (87 is the smallest number of pairs upon which any total correlation was based). If the fact that while 156 pairs are available for calculating

[1] It would probably have been better to ask for the numbers of *married* women between the ages 16–50, and we regret not to have done so (the illegitimate birth rate although high might perhaps have been neglected); but, for reasons which will appear, the correction is of less importance than would have been the case had we been working with English Registration Counties as units of grouping.

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

TABLE VII.

Partial Correlations: Population Constant.

Variables	Constants	Correlation
Birth Rate and Art. Feeding Rate	Poverty Rate and Infant Mort. Rate	– ·145 ± ·071
Birth Rate and Poverty Rate ...	Infant Mort. Rate and Art. Feeding Rate	·141 ± ·071
Birth Rate and Infant Mort. Rate	Poverty Rate and Art. Feeding Rate	·569 ± ·049
Poverty Rate and Art. Feeding Rate	Birth Rate and Infant Mort. Rate	·052 ± ·072
Poverty Rate and Infant Mort. Rate	Birth Rate and Art. Feeding Rate	·172 ± ·070
Art. Feeding and Infant Mort. Rate	Birth Rate and Poverty Rate ...	·662 ± ·041

the Birth-Rate—Infant Death Rate correlation—only 87 give the Poverty and Artificial Feeding Rates were due to chance, this might be regarded as a sound method. The ordinary process of deducing the probable error (see Pearson and Filon, Yule 1907 and Heron 1910) is not, however, directly applicable and the procedure adopted seems open to criticism. We have accordingly recalculated the whole of the coefficients using only the districts, 84 in number, for which each of the four variables is recorded. In Table VII A, the results of this work are given from which it will be seen that, while the correlations between the Infant Mortality Rate and the Rates of Artificial Feeding and Birth are not very seriously altered, the relative importance of the Poverty Rate, as indicated by its partial correlation with the Infant Death Rate, is sensibly different. This result, together with the evidence presented further on, makes it probable that some at least of the data are not random samples in the statistical sense and emphasises the necessity of caution in drawing conclusions.

We have now detailed the leading results of our inquiry but, before entering upon a consideration of them, several points need attention. As we stated in the introduction, uncertainty exists with respect to the value of the standards of measurement employed. While we have already noticed some objections and the answers to them, further evidence remains for consideration. This consists in a comparison of the results yielded by applying similar methods of analysis to other data and is probably the most valuable kind of check which can be used. Since few, if any, observers other than Groth and Hahn have tabled information respecting all four variables in a manner suitable to our method of study, the comparison can hardly be extended beyond the total coefficients; these however can be brought into relation with several other results which will now be noticed systematically.

TABLE VII A.

Correlations between Birth Rate (%$_0$ Females 16–50), Infant Mortality Rate, Poverty Rate and Artificial Feeding Rate in the 84 Rural Districts of Bavaria in which each Variable is recorded.

Variables	Correlation
Birth Rate (%$_0$ Females 16–50) and Artificial Feeding Rate	·461 ± ·058
Birth Rate (%$_0$ Females 16–50) and Poverty Rate	·413 ± ·061
Birth Rate (%$_0$ Females 16–50) and Infant Mortality Rate	·780 ± ·029
Poverty Rate and Artificial Feeding Rate	·356 ± ·064
Poverty Rate and Infant Mortality Rate	·526 ± ·053
Artificial Feeding Rate and Infant Mortality Rate	·727 ± ·035
Birth Rate (%$_0$ Females 16–50) and Population	·154 ± ·072
Poverty Rate and Population	·020 ± ·074
Infant Mortality Rate and Population	·265 ± ·068
Artificial Feeding Rate and Population	·251 ± ·069

Total Correlations with Constant Population.

Variables	Correlation
Birth Rate (%$_0$ Females 16–50) and Artificial Feeding Rate	·442 ± ·059
Birth Rate (%$_0$ Females 16–50) and Poverty Rate	·414 ± ·061
Birth Rate (%$_0$ Females 16–50) and Infant Mortality Rate	·776 ± ·029
Poverty Rate and Artificial Feeding Rate	·363 ± ·064
Poverty Rate and Infant Mortality Rate	·549 ± ·052
Artificial Feeding Rate and Infant Mortality Rate	·708 ± ·037

Partial Correlations with Constant Population.

Variables	Constants	Correlations
Birth Rate (%$_0$ Females 16–50) and Art. Feeding Rate	Poverty Rate and Infant Mort. Rate	– ·240 ± ·069
Birth Rate (%$_0$ Females 16–50) and Poverty Rate	Infant Mort. Rate and Artificial Feeding Rate	– ·016 ± ·074
Birth Rate (%$_0$ Females 16–50) and Infant Mort. Rate	Poverty Rate and Artificial Feeding Rate	·697 ± ·038
Poverty Rate and Art. Feed'g Rate	Birth Rate and Infant Mort. Rate	– ·035 ± ·074
Poverty Rate and Infant Mort. Rate	Birth Rate and Art. Feeding Rate	·320 ± ·066
Art. Feeding Rate and Infant Mort. Rate	Birth Rate and Poverty Rate	·622 ± ·045

The Correlation between the Birth Rate and the Infant Death Rate.

That a high birth rate tends to be associated with a high rate of infant mortality has long been recognised and we do not attempt to name all the writers who have commented upon the fact. In addition to the remarks of Malthus already quoted (see also Malthus, I. 469) we may refer to Wappäus (1859, I. 216; II. 321), Geissler (1885), Ganzen-Müller (1897), Mombert (1907), and Prinzing (1907). We shall restrict ourselves to a few of the papers containing material which can be analysed by the methods we have employed. The principal difficulty encountered in arriving at comparable results was the non-uniformity of the standard adopted in calculating birth rates. Most writers have contented themselves with expressing births in percentages of the whole population (crude birth rate); others base their ratio upon the number of wives at fertile ages or the number of women, married and single, at those ages. Newsholme and Stevenson (1906) have pointed out the errors which may result from a comparison of crude birth rates and have introduced a valuable process of correction similar to that universally adopted in the case of deaths. We are not indeed concerned with the birth rate as a measure of absolute fertility, but it is material to know whether the index employed has the same significance from district to district; if not, the coefficient of correlation will be misleading. If we determine the correlation between the birth rate and the infant mortality rate for English Registration Counties, the method of reckoning the former makes an enormous difference in the result[1]. Thus, taking the figures of Newsholme and Stevenson's paper, the correlation was ·81 when the crude birth rate was used, ·64 when the birth rate was expressed in terms of women between 15 and 45, ·53 if the rate was in terms of married women between 15–45 and only ·36 if Newsholme and Stevenson's standard birth rate was adopted. Now in the case of the Bavarian Rural Districts, the reduction of the coefficient effected when we passed from the crude birth rate to the rate in terms of all women between 16–50 was much less than for the English Registration Counties. It seemed to us that the explanation might be found in a more uniform distribution of sexes and ages in the rural districts of Bavaria than in the large and somewhat heterogeneous Registration Counties of England. As a matter of fact, we found that the coefficient of variation for the proportion of women (15–45) in the Registration Counties was 6·21; the same constant for the women (16–50) in the Bavarian Rural Districts was 3·79. The difference, 2·42 ± ·47, is statistically significant and goes far to explain the result; since, were the age, sex, and civil state proportionally the same in all districts, the correlation would be unaffected by the method of calculating the birth rate. For these reasons, while admitting the inferiority of our standard to that of Newsholme and Stevenson, we are inclined to doubt whether the correlation is in

[1] We have not corrected for absolute population except in the case of Bavaria.

reality far from the true value. We are disposed to believe that Newsholme's result (Newsholme, 1911; Local Government Board, 1910, p. 49), although large and significant may be too small, an English Registration County being, from many points of view, an unsatisfactory unit—for instance, it means the grouping together of towns and country districts, which we have shown to be undesirable[1].

TABLE VIII.

Correlation between Birth Rate and Infant Death Rate.

Source of Data	Coefficient of correlation	Computer	Remarks
27 Metropolitan Boroughs (Heron)	·50 ± ·10	Heron	Birth Rate based on married women (15–54).
17 Countries (Powys) ...	·93 ± ·02	Powys	Crude Birth Rate.
38 Cities and Towns of Massachusetts (State Board of Health) 1896	·42 ± ·09	Greenwood & Brown	,, ,, ,,
704 Rural Districts of the German Empire (Wurzburg)	·58 ± ·02	,, ,,	,, ,, ,,
161 BavarianRuralDistricts (Groth and Hahn)	·72 ± ·02	,, ,,	,, ,, ,,
156 ,, ,, ,,	·68 ± ·03	,, ,,	Birth Rate based on all women (16–50).
44 English Registration Counties (Newsholme and Stevenson) 1903	·36 ± ·09	Newsholme (p. 49)	Corrected Birth Rate (Newsholme and Stevenson).
44 ,, ,, ,,	·53 ± ·07	Greenwood & Brown	Birth Rate based on wives (15–45)
44 ,, ,, ,,	·64 ± ·06	,, ,,	Birth Rate based on all women (15–45).
44 ,, ,, ,,	·81 ± ·04	,, ,,	Crude Birth Rate.
44 ,, ,, (1901)	·86 ± ·03	,, ,,	,, ,, ,,
25 Swiss Cantons (Statistique de la Suisse)	·71 ± ·07	,, ,,	,, ,, ,,
143 Towns in Saxony (Ganzen-Müller)	·42 ± ·05	,, ,,	,, ,, ,,

In Table VIII we collect the results yielded by data obtained from various sources. The figures of Würzburg are particularly interesting because, in his elaborate and careful paper, not only did he separate the rural from the urban districts, but he dealt with the whole of the German Empire and thus provided sufficient material for us to consider the nature of the regression of Infant Mortality upon the Birth Rate. Proceeding in the ordinary way, we have $r = {\cdot}580$, $\eta = {\cdot}593$ and $\frac{\sqrt{n}}{.67449} \times \frac{1}{2}\sqrt{\zeta} = 2{\cdot}37$ (Pearson, Blakeman). Consequently the regression

[1] See on this Welton (Welton, 1911, p. 3).

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

may be regarded as effectively linear. This is of some importance, since our Bavarian material is not sufficiently extensive for us to be able to examine the regression in an altogether satisfactory way. The general impression which results from an examination of Table VIII is that our coefficient *may* be somewhat too large but that the difference from a reasonable expectation is not serious. Naturally, the coefficients in Table VIII measure total correlations, and do not bear comparison with our partial coefficients; this can be, *to some extent*, allowed for in the case of Heron's investigation, but we shall return to this shortly. In one set of observations which we have found, the poverty factor was experimentally eliminated. Thus Trap (cited by Saltet and Falkenburg, p. 15) published the result of an inquiry into the relation between size of family and the infant death rate, the parents being grouped under occupations. From these data (as given by Saltet and Falkenburg, p. 15) we obtained coefficients ranging from ·6 to ·9, but the material is not so presented as to admit of our determining the true probable errors, hence we do not feel justified in attaching importance to our deductions. It would have been possible to extend the figures in Table VIII, but sufficient have been given for our purpose; for a discussion of the relation between the birth rate and infant death rate, with other citations, reference should be made to Prinzing's treatise.

The Birth Rate and Poverty.

The statistical literature of this part of the subject is enormous and we shall observe the same limitations as before.

The most important contribution is Heron's monograph[1] (Heron, 1906). Heron dealt with the Metropolitan Boroughs and, as we have seen, founded the birth rate upon the number of wives aged 15–54 years. For his measure of poverty, Heron adopted several ingenious standards, the proportion of general labourers, the proportion of pawnbrokers, etc. All his coefficients are larger than ours but the difference is hardly material (Heron did not correct for populations in the way we have described but the effect of this does not seem great). Taking the result Heron obtained when he used the proportion of general labourers as a measure of poverty, it is to be remarked that the difference between his value and ours is less than its probable error (·07 ± ·11). Heron

[1] Niceforo (1911) has also studied the problem with the aid of the method of correlation, but the coefficient he used cannot be directly compared with those employed in this paper.

found that the correlation between the percentage of all kinds of pauperism and the birth rate was only ·199 ± ·145; we ourselves have even obtained negative values in the case of English Registration Counties. On the other hand, the Bavarian returns yield coefficients not dissimilar from those of Heron when he employed a more direct measure of poverty than we have been able to use. In view of all the circumstances mentioned and the fact that Heron considered an urban, we a rural, population, the results may be taken to show that the percentage of *Unterstützte* is a fairly reliable measure of poverty for the purposes of an inquiry like the present.

TABLE IX.

Correlation between the Birth Rate and Various Indices of Poverty.

Source of data	Nature of Birth Rate	Measure of Poverty	Correlation	Computer
27 Metropolitan Boroughs (Heron (1))	°/₀ of married women aged 15–54	General Labourers per 1000	·52 ± ·10	Heron.
27 Metropolitan Boroughs (Heron (1))	°/₀ of married women aged 15–54	Pawnbrokers, etc., per 1000	·62 ± ·08	Heron.
27 Metropolitan Boroughs (Heron (1))	°/₀ of married women aged 15–54	Proportion living more than 2 in a room	·70 ± ·07	Heron.
27 Metropolitan Boroughs (Heron (1))	°/₀ of married women aged 15–54	Proportion of children aged 10–14 employed	·66 ± ·07	Heron.
142 Bavarian Rural Districts (Groth and Hahn)	°/₀ of all women aged 16–50	°/₀ of publicly assisted persons	·44 ± ·05	Greenwood & Brown.
22 Districts of Hamburg (Mombert)	°/₀ of married women aged 15–45	°/₀ of dwellings rented at less than 300 marks	·88 ± ·03	Greenwood & Brown.
26 Districts of Leipzig (Mombert)	°/₀ of all women aged 15–40	°/₀ of dwellings rented at less than 250 marks	·83 ± ·04	Greenwood & Brown.
26 Districts of Leipzig (Mombert)	°/₀ of all women aged 15–40	°/₀ of dwellings with less than 2 heated rooms	·91 ± ·02	Greenwood & Brown.

In Table IX we collect Heron's and others' results for comparison. The general conclusion seems to be that although our crude coefficient may be too small, it would be rash, for the purposes of subsequent calculation, to increase its value upon conjectural grounds.

Poverty and the Rate of Infant Mortality.

The association between poverty and infant mortality is also a mere commonplace of the literature; *inter alios* may be named—Malthus (1817), Neumann (1893), Verrijn Stuart (1901), Rowntree (1902), Heron (1906), Newman (1906), Prinzing (1906), Saltet and Falkenburg (1907), Local Government Board Report (1910). Few of these writers, however, give material for expressing the association in a quantitative form;

Heron's memoir is an exception. From his Table A (Heron, 1906, p. 6) we find the correlation between poverty (as measured by the proportion of general labourers) and the Infant Mortality Rate to be ·40 ± ·11[1], a value which agrees admirably with our own findings. As we shall remark further on, it is doubtful how far this correlation really indicates an independent connection between the two variables, and we deeply regret being unable to submit our coefficient to other comparative tests. The conclusion to which our analysis tends on this point is, indeed, of considerable importance, *if true*; but we should like to have further warrant for believing it true than is afforded by a single comparison.

Artificial Feeding and the Rate of Infant Mortality.

Most of the authors already cited discuss this point; in addition we may refer to Creighton (1894, II. 766), Eröss (1903), The Sanitary Authority of Milan (1908), Seutemann (1909) and the Westminster Health Society (1910). In all the quantitative investigations we have found, the material is derived from a more or less complete census of the infant population of one city or borough, not from a series of districts as in Groth and Hahn's work. It was therefore necessary to adopt a

TABLE X.

Some Correlations between Survival and Breast-Feeding.

Source of Data	Correlation	Method	Notes
Barmen (Kriege and Seutemann, cited by Saltet and Falkenburg, p. 26*)	·41 ± ·01	Fourfold Table	The classification is into Breast- and Bottle-Fed children: correcting for income of Father raised r to ·44 ± ·01.
Westminster (City of Westminster Health Soc.)	·30 ± ·05	,,	†Partial Breast-Feeders included with full Breast-Feeders.
Derby (Howarth, cited by Newman, p. 238)	·33 ± ·02	,,	Mixed Feeders included with Breast-Feeders.
Milan (La Mortalità infantile in Milano)	·29 ± ·03	,,	Mixed Feeders included with Breast-Feeders.

* We have not found a copy of the Journal in which Kriege and Seutemann's paper appeared either at the British Museum or the two great Medical Libraries. There is a discrepancy between the figures as cited by Saltet and Falkenburg on p. 25 and p. 26 of their monograph. On p. 25, 4139 children are stated to have been alive on Aug. 15th, 1905. The table on p. 26 gives 4051. Possibly the difference is caused by omitting cases without particulars of Feeding. We have used the table on p. 26, *op. cit.*

† If the partial Breast-Feeders are included with the Bottle children, the correlation rises to ·69. A similar difference *but in the opposite direction* is produced in the Milan data by altering the classification.

[1] See p. 32.

different and hardly comparable method of measuring the correlation, the results of which appear in Table X. In our opinion, these coefficients are of but little use as comparative checks since, as noted in the table, the arithmetical value of the coefficient is greatly affected by the position of the dividing planes[1]. The most that can be said, and said with hesitation, is that the coefficients agree fairly well among themselves and perhaps suggest that our value (obtained be it noted by a different method and with a different arrangement) may be too high. As will be pointed out later on, there is some internal evidence for this belief; but we should have liked to find stronger extrinsic confirmation of our view.

General result of the comparative Tests.

It may, we think, be asserted that such portions of our work as could be checked by comparison with the results of other investigators using the same method, or by reducing other material in a similar fashion, do not exhibit any startling contradictions. At the same time, we fully recognise the limitations which have been imposed upon this process by the nature of the subject. At the least, nothing has been elicited which would tend to discredit conclusions based upon Groth and Hahn's material. We shall now proceed to a closer scrutiny of our own constants.

IV. *The nature of the total regressions and of the frequency distributions for the Bavarian data.*

The precise value to be assigned to a determination of the coefficient of correlation depends upon the nature of the distribution of which it is a constant. While r is always some measure of association, the value deduced in any particular case can only be directly compared with the result of another investigation, if there are reasonable grounds for thinking that the correlation surfaces are approximately similar. A failure to recognise this fact has led untrained users of the method of correlation into difficulties. The coefficient of correlation attains its maximum utility when the regression of one character upon another is linear. Since, if the frequencies are normal, the regressions must necessarily be linear, this point was first examined. From the nature of the material, no values below 0 or above 100 can occur, so that the frequencies could not be truly Gaussian, although normal curves might

[1] Doubtless the frequencies are not normal.

give sufficiently close approximations to them. The principal data used are given in Appendix I and it is apparent upon examination that the distribution of percentage artificial feeding has no resemblance at all to a normal curve (see diagram). The other three variables conform to the type of "cocked-hat" distributions, but, on fitting normal curves and testing with Elderton's (1902) tables, no close agreement could be obtained. The least bad result was with the Poverty Rates (P = ·04). Although the frequencies are not normal, it might still be that the regressions are linear and we have already found in the case of Würzburg's similar, but more extensive, data an example of effectively linear regression (Infant Death Rate upon Birth Rate). Owing to the relative sparsity of our data, the process employed in the case just referred to (Pearson, 1905; Blakeman, 1905) is not quite satisfactory, because the error introduced by grouping is considerable. For the purpose of determining η it was necessary to group the observations, but the value of r compared with it was deduced from the ungrouped figures. (The crude values, uncorrected for population, were used.)

TABLE XI.

Tests for Linearity of Regression (Crude Regressions).

Variables	r	η*	$\frac{\sqrt{N}}{\cdot 67449} \times \frac{1}{2}\sqrt{\eta^2 - r^2}$
Birth Rate and Poverty Rate	·426	·513 ·485	2·38 2·06
Infant Death Rate and Poverty Rate ...	·465	·538 ·534	2·41 2·34
Infant Death Rate and Artificial Feeding Rate	·761	·785 ·789	1·41 1·51
Infant Death Rate and Birth Rate	·462	·746 ·727	2·49 1·95
Poverty Rate and Artificial Feeding Rate ...	·369	·490 ·461	2·51 1·71
Birth Rate and Artificial Feeding Rate ...	·462	·581 ·521	2·23 1·91

* The two values correspond to the two regressions, that of x upon y and of y upon x.

It will be noticed in Table XI that the values in the third column never sensibly exceed 2·5; this fact, together with information to be derived from plotting the original data on a diagram, seems to suggest that the regressions are effectively linear. The conclusion to be drawn from these results is that a regression equation of the first degree connecting the four variables should be of interest. In order to obtain the constants of such an equation it is necessary to determine the standard deviations of the third order. As will be seen from the results collected in Appendix II, the means and standard deviations are somewhat affected by variations in the numbers of available pairs. For the present purpose we have adopted as the mean of each character that based upon all the available data and similarly the standard deviations of zero order were those deduced from the largest numbers of observations. We have not, in view of the relatively small effect on the coefficients of correlation which resulted from correcting for constant

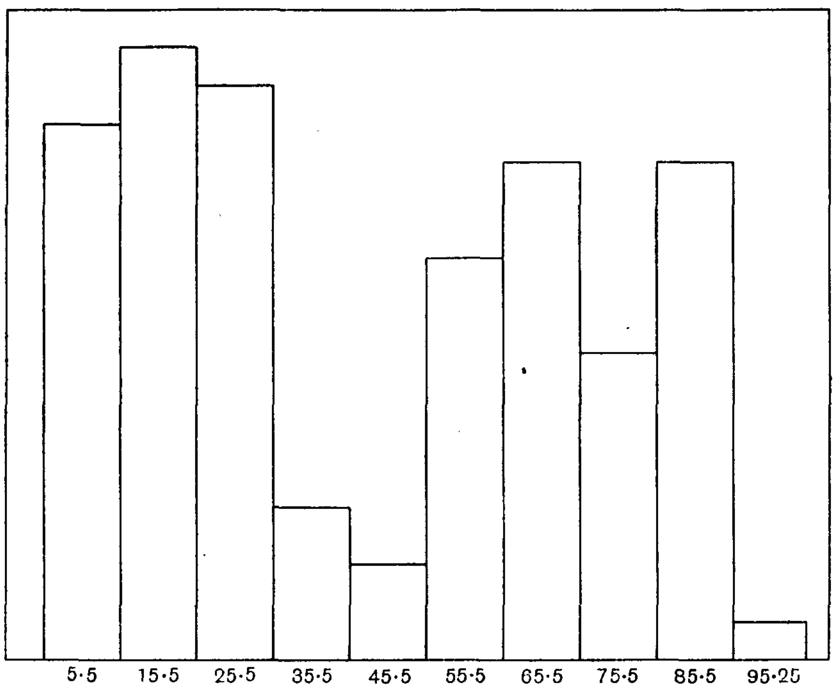

Fig. 1.

population, applied this process to the standard deviations. In determining the probable errors of the regression coefficients we have, as in the case of the partial correlation coefficients, given to n its smallest value. If x_1 be the rate of Infant Mortality, x_2 the Birth Rate (in terms of all women 16–50), x_3 the rate of Artificial Feeding and x_4 the Poverty Rate, then, measuring all rates from zero, we find:

$$x_1 = 1{\cdot}234\, x_2 + {\cdot}141\, x_3 + {\cdot}936\, x_4 - 4{\cdot}334.$$

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

Comparing the coefficients of regression with their probable errors, we have

$$1{\cdot}234 \pm {\cdot}129, \quad {\cdot}141 \pm {\cdot}012, \quad {\cdot}936 \pm {\cdot}387.$$

Turning to the 84 districts with complete records and, in this case, correcting the Standard Deviations for population, we obtained the equation:

$$x_1 = 1{\cdot}460\,x_2 + {\cdot}113\,x_3 + 1{\cdot}640\,x_4 - 8{\cdot}229.$$

The probable errors of the three regression coefficients are $\pm {\cdot}111$, $\pm {\cdot}011$, $\pm {\cdot}358$. As will be seen, there is a considerable difference between the weights attaching to the x_4 term in the two equations. With regard to the two other variables on the right-hand side, it will be noticed that a unit change in x_2 would be associated with a larger change in x_1 than would a variation of the same extent in x_3. This is, however, somewhat misleading from the practical standpoint since an absolute change of one per cent. in both the Birth Rate and the Artificial Feeding Rate would denote a much more substantial variation of the former than of the latter, the absolute values of their respective means being almost in the ratio of one to three. However, even if we allow for this fact, the Birth Rate term is of more weight than the Artificial Feeding term.

As will have been noticed in the diagram, the distribution of percentage artificial feeding tends to be bimodal, districts with percentages in the neighbourhood of fifty being in defect. This *may*, of course, indicate the real state of affairs. It may be that the mothers in one locality tend as a whole either to suckle their infants or to bring them up by hand. Fashion or some special economic conditions introducing heterogeneity might be responsible for such a result. On the other hand, it would not be right to neglect a different possibility. Groth and Hahn's data in this case are the result of circulating requests to the medical officers of the local vaccination departments (see p. 9) and they did not secure returns from anything like all the districts. It is just possible that when the medical officer found that either a very large or a very small proportion of the children brought under his notice were hand-fed, the circumstance attracted his attention and he did not fail to make a return; while in the instances in which the proportion of bottle-fed infants was only moderate, the information may have seemed commonplace and hardly worth recording. There is we imagine a somewhat general inclination to dwell upon extreme results in all inquiries, fostered by the tendency of writers to publish only the highest and lowest values of their census.

If anything of this kind has really occurred, the artificial feeding data are not a random sample but a selection. Pearson, in a classical memoir (1903), has shown that the value of a coefficient of correlation is much altered by selection. In this memoir, Pearson dealt with normal correlation and the case of the selected individuals being less variable than the population from which they were drawn, but has since indicated that the theorems then obtained were of greater generality than at first supposed. In the present case, the variability of the selection will be higher than that of the original population. In the case of normal correlation, it is easy to prove that a coefficient of correlation deduced by product moments from the selection will be larger than the value of r obtained by the same method from the original population. Illustrating this by an extreme case, let us suppose that part of the correlation surface which is included between parallel planes intersecting the surface at right angles to the horizontal plane and at distances $\pm h'$ from the centroid vertical has been removed. It will then be found that the coefficient of correlation deduced from what is left of the surface will be equal to $r \times \sqrt{\frac{1-p_2+2p_1}{1-p_2+2p_1 r^2}}$ where r is the coefficient of the whole surface,

$$p_1 = \frac{h}{\sqrt{2\pi}} \times e^{-\frac{1}{2}h^2}, \quad p_2 = \frac{1}{\sqrt{2\pi}}\left(\int_{-h}^{h} e^{-\frac{x^2}{2}} . dx\right), \quad h = \frac{h'}{\sigma_x}, \quad \sigma_x =$$

Standard Deviation of the "Selected" character.

For instance, if $h' = \sigma_x$ and $r = \cdot 5$, $p_1 = \cdot 2419$, $p_2 = \cdot 6827$, the apparent correlation is ·676. A similar expression can be deduced when the frequency near the centroid is not removed but diminished[1].

Since we have no adequate means of determining how far the preliminary speculation may be just, it is unnecessary to enlarge upon the mathematical consequences which might flow from it. We desire, however, to emphasize the fact that if any such process of selection have really been at work, the true correlation between Artificial Feeding and Infant Mortality Rates is lower than would appear from our analyses. Further statistics are necessary to settle this important point[2] and we sincerely hope that an inquiry comparable with, but more extensive than that of Groth and Hahn[3], may be embarked upon in our own country; evidently such a task is quite beyond the powers of private

[1] See Appendix III.

[2] Another possible source of bimodality which requires for its elucidation wider statistics is industrial differentiation.

[3] Compare also the difference between the coefficients of Tables VII and VII A.

individuals and could only be undertaken by the central government. In view of the very great and general interest which is aroused by the problems connected with infant mortality, such an inquiry ought not to be delayed, even although the amount of money and labour required were considerable.

V. *The Meaning of the Partial Correlations.*

Although the method of multiple correlation attains more nearly to the ideal of the scientific reasoner, who endeavours to isolate each member of a causal nexus, than do the ordinary processes of the official statistician, we must not over-estimate its powers. In particular, when the correlation is not normal, and in our case it is not normal, the partials represent in general no more than an average measure of association (*vid.* Yule, 1911, p. 248). We must also bear in mind the various objections to and limitations of our material which have been dwelt upon in the course of this paper.

At the same time, the results deduced seem of sufficient interest and importance for it to be worth while devoting some paragraphs to a discussion of their possible meaning. We hope that, in considering these speculations, the reader will bear in mind our previous criticisms and will understand that our conclusions are put forward mainly as suggestions which can be of service to future inquirers.

Infant Mortality and the Birth Rate.

The correlation between the Birth Rate and the Rate of Infant Mortality is not a new discovery and has, we think, been usually explained along the following lines. Those persons who marry young and do not restrain their procreative impulses tend to be drawn from a class of society in which the instinct of commercial prudence is not highly developed and the economic environment is bad. As the family increases in size, the environment becomes less and less satisfactory, with the growing number of mouths to feed. In other words, the association between the birth rate and the infant mortality rate is a consequence of the association between the former and poverty. Now, if our method had been completely adequate to the task of keeping the poverty factor constant, had our measure of economic status been thoroughly satisfactory, this current explanation would have to be modified in view of the results we have obtained. While we cannot pretend to have fully satisfied these conditions, enough has been done to render it, in our judgment, probable that the association may be more direct than usually thought. A tempting speculation is suggested by the view entertained by the English eugenists (see Schuster, 1905; Pearson, 1907 and 1909; Heron, 1907, and many other publications by Professor Pearson and his pupils) that stocks differentiated by physical or mental stigmata of inferiority are more fertile than the average. This phenomenon may, we think, be largely attributed to the employment of preventive checks by the superior artisan and middle class populations. It is not probable that successful or dominant types are naturally infertile; indeed it has been shown that longevity and fertility are correlated (Beeton, Pearson and Yule, 1901), while there is much evidence that longevity is an heritable character (Beeton and Pearson, 1899 and 1901). The relation between fertility and longevity was, however, determined on material probably little affected by artificial restrictions of fertility (see Beeton, Pearson and Yule, *op. cit.* pp. 160–1). The results just cited suggest that there is no inevitable association between fertility and inferiority and compel us to refer the fact—assuming it to be a fact—to the voluntary measures at present so much debated. The causation of the result does not, however, affect the argument based upon it. If the facts be as stated, it is easy to interpret our correlation. It can hardly be doubted that the offspring of pathologically weighted stocks will be, in general, less resistant to the ordinary dangers of childhood than members of normal families. The correlation between the birth rate and the infant death rate should be interpreted as a consequence of the association between the former and natural inferiority. A necessary inference would be that, under modern conditions, the birth rate is itself a better criterion of unfitness than any conventional measure of poverty such as the one here used. To avoid the risk of misconception, fostered by the pernicious habit of quoting detached phrases from a scientific paper, we once more emphasize the words *under modern conditions.* If those who are worthy citizens are relatively infertile, it is not in virtue of any immutable decree of fate but a consequence of conditions which may not have existed in the past (see Heron, 1906), are possibly not universal now and will *perhaps* disappear in the future. We are not discussing what must be but what is or—to be quite accurate—what may be. Another conclusion to be drawn from our present hypothesis is that the infant death rate must be, to some extent, selective; to what extent, is another question.

It is to be remarked that the whole of the previous argument falls to the ground if it be not a fact that the gross fertility of pathologically weighted stocks is greater than the average. Pearson and Heron's results have been recently impugned by Weinberg (1911) but, as indicated in the footnote, this author's criticism appears to us to be based upon an arbitrary assumption[1].

At the same time, we are not prepared to deny that the relationship between fertility and mental or physical unfitness deserves further study.

Whatever may be thought of the speculation outlined above—we present it with the utmost diffidence—one conclusion appears highly probable. It is that the association between the actual Birth Rate and the Infant Death Rate is of considerable and direct importance[2], a phenomenon not lightly to be put on one side as of little immediate interest. We state this with some confidence, although a directly opposite opinion has received the sanction of so high an authority as Newsholme (1911, p. 44). Newsholme himself found a substantial correlation between the two rates (see p. 19) and we believe that, for the purpose in hand, the Bavarian data are more satisfactory than the returns from English Registration Counties. In working with the latter, it is impossible to avoid the difficulties introduced by mixing the populations of urban and rural districts. Other variables (for instance the effect of varying industrial conditions) are introduced into a problem already sufficiently complex, with the result that it becomes difficult to interpret the statistical constants deduced. It is needless to say that our conclusions on this as upon other points are subject to revision in the light of wider experience.

[1] Weinberg asserts that the difference in fertility between normal and pathological stocks found by Pearson, is spurious and due to the fertility of the normal stocks having been determined from the parental, that of the pathological stocks from the filial generation. He publishes (*op. cit.* p. 690) a so-called "General Proof" that the mean size of family *based upon the "Angaben" of the children,* will be greater than that deduced from a count of the parents. Obviously, if we have, say, two families, one with one, the other with two children, the mean size of family, based on the parental "Angaben" is 1·5, while if the children are questioned, the "Angabe"—"Two in Family"—will be given twice and "One in Family" once. From these replies, the mean family might be reckoned as 1·67. This appears to be all that Weinberg has succeeded in proving, and it seems to us unwarrantable in him to suppose that Pearson and Heron adopted so singular a method of reckoning. Evidently, if the names and pedigrees of the children are known—in the cases singled out for criticism by Weinberg they were known—no one could fall into this trap. We accordingly dissent from the conclusion stated by Weinberg at the foot of p. 690 (*op. cit.*) If Weinberg's "General Proof" has an interpretation other than that placed upon it in this note, we tender our apologies to him for having failed to discover it.

[2] See also the important discussion in Westergaard's treatise (p. 368).

Artificial feeding and the infant death rate.

Here again, *in so far as we have succeeded in neutralising the factor of social environment*, we have got rid of one obvious cause of infant mortality associated with artificial feeding, viz., the increased liability to septic contamination of the feeding bottles in poor households and the use of an inferior milk. But, apart from these accessory conditions, it is not surprising to find a positive association. Numerous experiments have demonstrated that the milk of each species is to some extent peculiar to it. After all, it does not seem likely that anything devised as a substitute for a normal secretion will on the average—neglecting the absolutely large but relatively small number of instances in which the natural secretion is demonstrably inadequate or of poor quality—vie with human milk as a food for infants. Further, it is not very far-fetched to suppose that all the reproductive functions are intimately correlated in a physiological sense, that diminished activity of the mammary glands may be associated with inferior viability of the off-spring. It is, however, clear that the significant correlation between infant mortality and the prevalence of artificial feeding is the most hopeful result of our inquiry. While it may not be, and probably will not be, possible to arrive at a thoroughly efficient substitute for the method of breast feeding, it is undoubted that the progress of physiological chemistry and domestic hygiene will, step by step, enable us to approximate to the ideal. But, for reasons already given at length, we cannot think that the factor of artificial feeding is equal in importance to that measured by the Birth Rate.

Infant mortality and poverty.

Since we have, so far as our data permitted, eliminated the influence of the Birth Rate and Method of Feeding, it would be expected that the correlation might undergo some change, but the actual result in this, as well as the other correlations into which the factor of poverty enters, is at first sight perplexing. It is desirable to preface our discussion of results involving poverty with a few words of caution to the reader unversed in statistical methods. In interpreting the first regression equations, we laid no stress upon the term involving the poverty rate because the probable error of its constant was relatively large in one case. This was justifiable, but we should go beyond our evidence were we to assert that a change in the poverty rate is definitely uncorrelated with a change in the rate of mortality. All we can say is, that, from the data before us, we have been unable to demonstrate that the given measure of poverty is closely associated with the Infant Mortality Rate (cf. the three partial coefficients involving the poverty rate with their probable errors as entered in Table VII)[1]. A negative conclusion must not be transformed into a positive assertion. Let us now compare our partial coefficients with those to be deduced from Heron's (1906) memoir. Denoting the Birth Rate, the Artificial Feeding Rate, the Poverty Rate, and the Infant Death Rate by the subscripts, 1, 2, 3, 4, using Yule's notation for the partial coefficients and taking the Proportion of General Labourers as a measure of the Poverty Rate for Heron's data, we obtain the following comparative results.

	Heron	Greenwood & Brown
$r_{13 \cdot 4}$	·40 ± ·11	·14 ± ·07
$r_{13 \cdot 24}$	—	·14 ± ·07
$r_{14 \cdot 3}$	·37 ± ·11	·64 ± ·04
$r_{14 \cdot 23}$	—	·57 ± ·05
$r_{34 \cdot 1}$	·19 ± ·13	·28 ± ·07
$r_{34 \cdot 12}$	—	·17 ± ·07

If we compare our second order coefficients with the first order coefficients of Heron's work (his data did not enable him to deal with a fourth variable), we find the differences are, + ·26 ± ·13, − ·20 ± ·12, ·02 ± ·15. While it cannot be said that any one of these differences is necessarily significant, it is worthy of remark that our Birth Rate-Poverty Rate coefficient is smaller, our Birth Rate-Infant Death Rate coefficient larger than Heron's. We are tempted to explain the second finding by the inferiority of our method of measuring the Birth Rate; and we have little doubt that his more satisfactory measure of poverty has been influential in the first case. Still, when all allowances are made, we should be disposed to attribute greater importance to the fact that we are not only dealing with another country, but also with a different class of population. Heron pointed out—and it is not the least interesting part of a valuable paper—that the intensities of some of the correlations he found were very different at different epochs[2]. Another point to be noticed is that he obtained a higher correlation

[1] Note also the relatively and absolutely greater weight attaching to the poverty term in the equation deduced from the 84 districts of Table VII A.

[2] In the London Boroughs, the correlation between Infant Mortality Rate and Birth Rate was *negative* in 1851, − ·30 ± ·10, positive in 1901, + ·51 ± ·10. (Heron, 1906, p. 18.)

between the Poverty Rate and the Birth Rate when the proportion of pawnbrokers was used as the measure of poverty than when the proportion of general labourers was chosen, although the difference may not be significant. The suggestion immediately arises as to whether the sinister associations of poverty in urban districts may really be, not functions of pure poverty, regarded as an environmental condition, but of a peculiar type of organisation.

In life as we know it in a modern city, the economic struggle for existence is so keen, the penalty of failure so heavy, that the weakest must be forced to the wall. Whether success in this struggle be the highest kind of success is an irrelevant consideration; at least all those who, under any definition of a question-begging term, are "fit" will endeavour not to fail. It accordingly results that, in towns, economic success is, to some extent, a measure of civic fitness. In the case of rural districts, however, the economic test, although still a rough approximation to the truth, may be definitely less adequate than in industrial towns. Were the preceding remarks substantiated, our results would be explained, but we have no evidence before us of a character sufficiently definite to justify further speculations[1].

The Birth Rate and the Artificial Feeding Rate.

As previously remarked (p. 10), Groth and Hahn concluded that prolonged lactation tended markedly to lower the Birth Rate, a somewhat wide-spread belief in certain classes. The fact that the total correlation between the Artificial Feeding and Birth Rates is positive would tend to support the conclusion, but the effect of neutralising other factors by the method of partial correlation is such that we are inclined to doubt its truth. If our view be correct, this case shows the importance of employing the language of multiple correlation in problems such as the one before us.

VI. *Conclusions respecting the chief factors of Infant Mortality.*

We have now described to the best of our ability both the good and bad points of the material collected by Groth and Hahn, together with the methods we adopted in analysing it. With respect to Infant Mortality in the rural districts of Bavaria, we conclude:

(1) Of the factors measured by Groth and Hahn the most important is the Birth Rate. A high Birth Rate tends to be associated

[1] It is impossible to pass from speculations to well-weighed conclusions until we have a stricter definition and better criterion of poverty than at present available.

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

with a high Infant Death Rate and this association cannot be explained by any inter-relations between either variable and proportional poverty or artificial feeding.

(2) It follows, as a corollary to (1), that a considerable share in the causation of infant mortality should be attributed to a factor beyond the ordinary sphere of preventive medicine.

(3) The habit of artificially feeding infants is definitely correlated with their rate of mortality, so that some improvement in the latter may be reasonably expected with the growth of physiological knowledge and public hygiene.

(4) It has not been possible to demonstrate any unambiguous association between poverty, as measured by Groth and Hahn, and the death rate of infants.

(5) The same remark applies to the relation between breast feeding and the Birth Rate, as in the case of Poverty and the Infant Death Rate.

VII. *The selective value of a high rate of Infant Mortality.*

As mentioned in the introduction to this paper, much attention has been devoted to the supposed selective importance of a high rate of infant mortality. The two most recent investigators are Yule (Local Gov. Board, 1910, Appendix I) and Snow (1911) and their results are, to some extent, discordant. The difficulty of the problem is very considerable. The fact that a high rate of infant mortality is usually associated with a high death rate at all ages, does not prove the non-selective character of the former, since the influence of an unfavourable environment might well act in the same sense at all ages. The efforts of both workers have accordingly been directed to discover an adequate measure of this environment and then, having neutralised it by the method of multiple correlation, to measure the partial correlation between the mortalities of infancy and later ages. We have grave doubts as to whether the problem can be solved with the help of mortality statistics alone and we are unable to share the confidence with which Snow appears to regard certain of his conclusions. It is not, however, necessary for us to enter upon a critical examination of a subject of great intricacy since, although Groth and Hahn touch upon the point, they do not provide data of an altogether satisfactory kind. What they deal with is the relation between Infant Mortality, the percentage of approved recruits, the prevalence of Artificial Feeding and the Death Rate of young adults. Unfortunately they do not bring into relation the returns of Infant Mortality for, say, 1880 and the recruiting statistics for 1900.

While it is quite likely that inter-local differences at any one time are much greater than the changes within any one district over a small number of years, we can have but little confidence in the data for the present object. On these grounds, we have only calculated one set of constants, taking as variables, Infant Mortality, Artificial Feeding, Mortality at ages 11–20 and the Fitness of Recruits.

TABLE XII.

Correlation between Infant Mortality, Death Rate (11–20 years), Percentage Fitness of Recruits and Artificial Feeding, Bavarian Rural Districts.

Variables	Observations	Correlation
Death Rate (11–20) and Infant Mortality Rate ...	92	·084 ± ·070
Infant Mort. Rate and Percentage Fitness of Recruits	92	·125 ± ·069
Death Rate (11–20) and Percentage Fitness of Recruits	92	– ·195 ± ·068
Death Rate (11–20) and Artificial Feeding Rate ...	92	·159 ± ·069
Art. Feeding Rate and Infant Mort. Rate	92	·755 ± ·030
Art. Feeding Rate and Percentage Fitness of Recruits	92	– ·054 ± ·070

Correlation between each of the Variables and Population.

Death Rate (11–20) and Population	– ·072 ± ·070
Infant Mortality and Population	– ·061 ± ·070
Percentage Fitness of Recruits and Population	– ·055 ± ·070
Artificial Feeding and Population	– ·055 ± ·070

Total Correlation with Population Constant.

Death Rate (11–20) and Infant Mortality Rate	·080 ± ·070
Infant Mortality Rate and Percentage Fitness of Recruits ...	·122 ± ·069
Death Rate (11–20) and Percentage Fitness of Recruits	– ·200 ± ·068
Death Rate (11–20) and Artificial Feeding Rate	·155 ± ·069
Artificial Feeding Rate and Infant Mortality Rate	·754 ± ·030
Artificial Feeding Rate and Percentage Fitness of Recruits ...	·057 ± ·070

Partial Correlations with Population Constant.

Variables	Constants	Correlation
Death Rate (11–20) and Inf. Mort. Rate	Percentage Fitness of Recruits and Art. Feeding Rate	– ·009 ± ·070
Inf. Mort. Rate and Percentage Fitness of Recruits	Death Rate (11–20) and Art. Feeding Rate	·245 ± ·066
Death Rate (11–20) and Percentage Fitness of Recruits	Inf. Mort. Rate and Art. Feeding Rate	– ·186 ± ·068
Art. Feeding Rate and Death Rate (11–20)	Inf. Mort. Rate and Percentage Fitness of Recruits	·102 ± ·070
Art. Feeding Rate and Inf. Mort. Rate	Death Rate (11–20) and Percentage Fitness of Recruits	·765 ± ·029
Art. Feeding Rate and Percentage Fitness of Recruits	Death Rate (11–20) and Inf. Mort. Rate	– ·204 ± ·067

The results appear in Table XII[1].

On account of the objections already mentioned, we do not propose to discuss the table at length. Apart from the Artificial Feeding-Infant Mortality correlation, already considered in detail, only two of the coefficients exceed thrice the probable error. If the percentage fitness of recruits be accepted as a measure of fitness in the true sense—as to which we have grave doubts—then the result affords some evidence of the selective value of an infant death rate. If, on the other hand, we regard the recruiting standard as a measure of environment and the death rate at ages 11–20 as a test of fitness, no statement can be made. For our own part, we should hesitate to draw any definite conclusions whatever.

Concluding Remarks.

Any impartial reader of this memoir will, we think, be apt to form some such conclusion as the following.

Since we have thought it desirable to devote a good deal of labour to the reduction of Groth and Hahn's material, we are not likely to have underestimated its value. Since, in the face of so natural a prepossession in favour of the evidence, we have been compelled again and again to emphasize the difficulty of basing trustworthy conclusions upon our analytical results, since, lastly, these results conflict with opinions often and confidently expressed, our investigation may be regarded as of no value.

We are far from suggesting that such an opinion is without foundation, but we believe some important submissions can be made in arrest of judgment. We would inquire whether the statements upon which popular and current medical opinions depend do really deserve more credence than the data here examined; whether much that passes for evidence is really more than optimistic dogma.

A knowledge of the factors controlling the rate of infant mortality can, we believe, only be attained by the methods of statistical science, indeed only by a refined handling of those methods. The day is far distant when the problem can safely be discussed by the proverbial "man in the street." We are not, therefore, without hope that the present memoir, in spite of—perhaps, because of—its tentative and incomplete character, will be of service to fellow-labourers.

[1] For the purposes of this table only those data were used which recorded all four variables. This accounts for the slight discrepancy between the value of the Artificial Feeding-Infant Mortality coefficient of this table and of those given previously.

REFERENCES.

BEETON, M. and PEARSON, K. (1899). A First Study of Inheritance of Longevity and the Selective Death-Rate. *Proc. Roy. Soc.* LXV. 290.

BEETON, M. and PEARSON, K. (1901). Inheritance of the Duration of Life and the Intensity of Natural Selection in Man. *Biometrika*, I. 50.

BEETON, M., PEARSON, K., and YULE, G. U. (1901). On the Correlation between Duration of Life and the Number of Offspring. *Proc. Roy. Soc.* LXXVII. 159.

BLAKEMAN, J. (1905). On Tests for Linearity of Regression in Frequency Distributions. *Biometrika*, IV. 332.

CREIGHTON, C. (1895). *A History of Epidemics in Britain.* Vol. II. Cambridge.

ELDERTON, W. PALIN (1902). Tables for Testing the Goodness of Fit of Theory to Observation. *Biometrika*, I. 155.

ERÖSS, J. (1895). Sterblichkeitsverhältnisse d. Neugeborenen u. Säuglinge. *Zeitschr. f. Hygiene u. Infectionskr.* XIX. 371.

GANZEN-MÜLLER, K. (1897). Die Fruchtbarkeits und Sterblichkeitsverhältnisse in sämtlichen Städten Sachsens 1891–5. *Zeitschr. d. Sächs. Stat. Bur.* XLIII. 142.

GEISSLER, A. (1885). Ueber den Einfluss d. Säuglingssterblichkeit auf die eheliche Fruchtbarkeit etc. *Zeitschr. f. Sächs. Stat. Bur.* p. 23.

GROTH and HAHN (1910). *Die Säuglingsverhältnisse in Bayern.* München (Lindauer) (*Sonderabdr. a. d. Zeitschr. d. K. Bayer. Stat. Landesamts.* Jahrg. 1910).

HERON, D. (1906). *On the Relation of Fertility in Man to Social Status etc.* London (Dulau).

HERON, D. (1907). *A First Study of the Statistics of Insanity etc.* London (Dulau).

HERON, D. (1910). On the Probable Error of a Partial Correlation Coefficient. *Biometrika*, VII. 411.

LOCAL GOVERNMENT BOARD (1910). *Supplement to the Report of the Board's Medical Officer.* Cd. 5263.

MALTHUS, T. R. (1817). *An Essay on the Principle of Population.* 5th Edition, London.

MASSACHUSETTS (1896). *Report of State Board of Health.*

MILAN (1908). *La Mortalita infantile in Milano.*

MOMBERT, P. (1907). *Studien zur Bevölkerungsbewegung in Deutschland.* Karlsruhe (Braun).

NEUMANN, F. J. (1893). Zur Lehre von den Lohnengesetzen (4) Pauperismus u. Kindersterblichkeit. *Jahrb. f. Nationalökon.* 3rd Series, V. 617.

NEWMAN, G. (1906). *Infant Mortality, A Social Problem.* London.

NEWSHOLME, A. (1911). *The Declining Birth-rate, Its National and International Significance.* London (Cassell).

NEWSHOLME, A. and STEVENSON, T. H. C. (1906). The Decline of Human Fertility in the United Kingdom and Other Countries as shown by Corrected Birth Rates. *Journ. Roy. Stat. Soc.* LXIX.

NICEFORO, A. (1911). Contribution à l'étude des Correlations entre le Bienêtre économique et quelques Faits de la Vie démographique. *Journ. d. l. Soc. d. Stat. d. Paris*, p. 322.

PEARSON, K. (1903). On the Influence of Natural Selection on the Variability and Correlation of Organs. *Phil. Trans.* A. CC. 1.

PEARSON, K. (1905). *On the General Theory of Skew Correlation and Non-Linear Regression.* London (Dulau).

PEARSON, K. (1907). *A First Study of the Statistics of Pulmonary Tuberculosis.* London (Dulau).

PEARSON, K. (1909). *The Problem of Practical Eugenics.* London (Dulau).

PEARSON, K. and FILON, L. N. G. (1898). On the Probable Errors of Frequency Constants. *Phil. Trans.* A. CXCI. 229.

POWYS, A. O. (1901 and 1905). Data for the Problem of Evolution in Man. *Biometrika*, I. 30 and IV. 233.

PRINZING, F. (1906). *Handbuch der Medizinischen Statistik.* Jena (Fischer).

ROWNTREE, R. S. (1902). *Poverty: a Study of Town Life.* London.

SALTET, R. H. and FALKENBURG, P. H. (1907). *Kindersterblichkeit besonders in den Niederlanden.* Amsterdam (Müller).

SCHUSTER, E. (1905). Hereditary Deafness. *Biometrika*, IV. 465.

SEUTEMANN, K. (1909). Oeffentliche Säuglingsfürsorge und Sterblichkeitsstatistik. *Jahrb. f. Nationalökonom. u. Stat.* XXXVII. 87.

SILBERGLEIT, H. (1897). Kindersterblichkeit in Europ. Grossstädten. *Arb. d.* VIII. *Int. Cong. f. Hyg. u. Demog.* Budapest.

SNOW, E. C. (1911). *The Intensity of Natural Selection in Man.* London (Dulau).

STERBETAFELN (1910). *Deutsche Sterbetafeln f. das Jahrzehnt* 1891 *bis* 1900. Berlin.

STUART, VERRIJN (1901). Untersuch. ü. d. Bezieh. Wohlstand etc. und Kindersterb. in den Nierderlanden. *Zeitschr. f. Sozialwissen.* IV. 649.

Suisse, Statistique de la (1901). CXXVIII. 35.

VOGL, A. VON (1909). *Die Sterblichkeit der Säuglinge in ihrem Territorialen Verhalten etc.* München, 1909 (Lehmann).

WÄPPAUS, J. E. (1859–61). *Allgemeine Bevölkerungsstatistik.* Leipzig.

WEINBERG, W. (1910). Die Rassenhyg. Bedeutung der Fruchtbarkeit. *Archiv f. Rassenhyg. u. Gesellschaftsbiol.* VII. 684.

WELTON, T. A. (1911). *England's Recent Progress.* London (Chapman and Hall).

WERNICKE, J. (1889). *Das Verhältnis z. Geborenen u. Gestorbenen in Hist. Entwickelung.* Jena (Fischer).

WESTERGAARD, H. (1901). *Die Lehre von der Mortalität und Morbidität.* Jena (Fischer).

WESTMINSTER HEALTH SOCIETY (1910). *Sixth Annual Report of the City of Westminster Health Society.* London (Riordan).

WÜRZBURG, A. (1887). Die Säuglingssterblichkeit im Deutschen Reiche etc. *Arbeit. a. d. Kaiser. Gesundheitsamte*, II. 208 and 343.

YULE, G. U. (1907). The Theory of Normal Correlation for any Number of Variables, etc. *Proc. Roy. Soc.* A. LXXIX. 182.

YULE, G. U. (1911). *An Introduction to the Theory of Statistics.* London (Griffin).

APPENDIX I.

Principal Data analysed in the text (Columns (2) and (4) are from unpublished data supplied by the Bavarian Statistical Office, the remainder are from Groth and Hahn's work).

Name of District	1 Population average, 1900—4	2 Number of women aged 16—50, Census 1901	3 Crude Birth Rate, 1900—4	4 Birth Rate in terms of Column 2	5 Poverty Rate, 1899—1902	6 Infant Death Rate, 1900—4	7 Percentage of children not Breast-fed, 1903—8	8 Number of children upon which (7) is based
Oberbayern:								
Aibling	22512	5093	3·46	15·27	2·51	24·6	63·5	1159
Aichach ...	27437	6115	4·26	19·09	2·80	33·6	63·3	770
Altötting ...	35177	8044	3·35	14·65	3·15	25·6	83·8	1122
Berchtesgaden ...	21512	5172	3·18	13·21	3·98	20·0	82·4	437
Bruck	26013	5440	4·17	19·94	1·51	38·6	—	—
Dachau	26572	5841	4·40	20·01	1·80	34·7	70·7	792
Ebersberg ...	25953	5569	3·81	17·77	2·18	35·3	—	—
Erding	41944	8953	3·88	18·18	3·09	34·8	—	—
Freising... ...	33369	7475	4·54	20·26	2·98	34·3	71·8	501
Friedberg ...	34425	8041	4·73	20·23	2·68	39·7	68·0	1053
Garmisch ...	13496	3072	3·15	13·85	1·47	21·0	—	—
Ingolstadt ...	24627	5227	5·08	23·94	2·38	40·7	—	—
Landsberg ...	24828	5290	3·85	18·06	1·80	28·3	—	—
Laufen	32954	7289	3·29	14·89	2·49	22·1	68·5	833
Miesbach ...	34834	7683	3·54	16·07	1·32	22·0	84·2	228
Mühldorf ...	37968	8401	3·75	16·93	2·35	29·8	87·2	476
München I ...	44092	10214	4·40	19·00	1·49	38·4	75·4	999
München II:								
†Starnberg ...	18782	7924	3·60	15·26	1·91	25·6	—	—
Wolfratshausen...	17166		3·11		2·73	26·7	90·1	462
Pfaffenhofen ...	35599	7468	4·58	21·83	2·95	38·5	70·2	991
Rosenheim ...	38248	8541	3·44	15·40	1·84	21·6	55·0	1545
Schongau ...	20453	4384	3·56	16·60	1·89	24·7	—	—
Schrobenhausen ...	20414	4498	4·52	20·52	2·99	35·7	58·8	583
Tölz	16834	3898	2·98	12·87	2·83	22·2	77·0	400
Traunstein ...	41589	9316	3·39	15·11	1·92	22·1	69·8	696
Wasserburg ...	36908	7939	3·38	15·69	3·12	27·8	56·4	553
Weilheim ...	31637	6960	3·58	16·27	2·17	22·2	—	—
Niederbayern:								
Bogen	31631	6824	4·29	19·88	4·18	33·4	70·8	364
Deggendorf ...	38274	8228	3·99	18·56	4·00	32·3	—	—
Dingolfing ...	22260	5078	3·86	16·91	3·65	33·9	86·7	581
Eggenfelden ...	37419	8244	3·88	17·59	2·09	32·2	91·9	333
Grafenau ...	18885	3928	4·37	21·02	5·43	26·2	—	—
Griesbach ...	33756	7353	3·75	17·23	3·08	28·4	82·8	1224
Kelheim ...	33794	7422	4·52	20·57	3·34	41·2	—	—
Kötzting ...	25787	5505	4·53	21·21	3·87	25·6	—	—
Landau a. I. ...	22995	4992	3·98	18·33	3·46	31·4	—	—
Landshut ...	29216	6409	4·23	19·28	3·01	33·7	—	—

† The number of women aged 16–50 only obtained for the joint district in these cases.

Name of District	1 Population average 1900—4	2 Number of women aged 16—50, Census 1901	3 Crude Birth Rate, 1900—4	4 Birth Rate in terms of Column 2	5 Poverty Rate, 1899—1902	6 Infant Death Rate, 1900—4	7 Percentage of children not Breast-fed, 1903—8	8 Number of children upon which (7) is based
Niederbayern (*cont.*):								
Mallersdorf ...	23529	5357	4·24	18·61	3·37	36·2	77·4	1297
Passau	42367	9381	3·76	16·98	3·65	29·4	—	—
Pfarrkirchen ...	36666	8318	3·75	16·52	2·86	30·3	90·4	466
Regen	27603	5883	4·57	21·44	3·45	29·6	39·1	884
†Rottenburg ...	18084	7546	4·35	20·76	3·92	34·1	80·8	468
Mainburg ...	16735		4·67		2·99	37·2	75·9	552
Straubing ...	22255	5097	3·92	17·13	2·90	35·2	83·8	1156
Viechtach ...	22492	4868	4·27	19·75	3·49	27·3	—	—
Vilsbiburg ...	30733	6906	3·80	16·91	2·50	31·8	—	—
Vilshofen ...	42948	9389	3·96	18·11	4·61	31·9	—	—
Wegscheid ...	17268	3722	3·67	17·05	4·43	23·8	—	—
Wolfstein ...	30136	6402	4·00	18·81	3·12	27·2	50·5	376
*Pfalz (Bavarian Palatinate):								
Bergzabern ...	38683	8776	3·15	13·87	—	14·6	16·4	456
†Dürkheim ...	28954	19475	2·75	13·16	—	16·6	28·1	595
Neustadt a. H. ...	51263		3·45			16·7	12·5	939
Frankenthal ...	62563	14256	3·67	16·12	—	18·1	24·8	749
Germersheim ...	53979	12098	3·36	14·99	—	19·3	—	—
Homburg ...	65052	14015	4·18	19·42	—	13·3	—	—
Kaiserslautern ...	84975	19913	3·73	15·90	—	14·2	15·0	1429
Kirchheimbolanden	26238	6035	3·22	13·99	—	16·3	16·4	664
Kusel	44718	9860	3·45	15·65	—	10·9	12·1	347
Landau	70769	16655	2·95	12·55	—	17·5	—	—
Ludwigshafen a. Rh.	97010	21725	4·65	21·77	—	22·2	—	—
Pirmasens ...	74531	16451	4·27	19·34	—	17·3	6·8	427
Rockenhausen ...	38685	8831	3·14	12·57	—	12·0	6·8	250
Speyer	39307	9652	3·81	15·53	—	25·0	—	—
†St Ingbert ...	38459	17977	4·40	17·72	—	13·7	6·1	668
Zweibrücken ...	43289		3·44			14·5	—	—
Oberpfalz:								
Amberg	25679	5893	4·17	18·16	1·81	29·3	—	—
Beilngries ...	28853	6536	4·19	18·48	2·74	38·7	—	—
Burglengenfeld ...	27312	5934	4·49	20·65	2·82	35·5	—	—
Cham	28925	6058	4·38	20·93	2·96	27·3	25·0	372
Eschenbach ...	22349	5150	3·32	14·42	2·36	18·4	—	—
Kemnath ...	22901	4820	3·62	17·19	2·58	15·5	—	—
Nabburg ...	17948	4068	3·92	17·29	3·45	26·0	23·4	516
Neumarkt ...	23668	7357	3·79	14·78	3·14	33·6	25·4	421
Neunburg v. W. ...	15119	3289	3·96	18·21	1·73	27·5	—	—
Neustadt a. W.-N.	36620	7980	3·82	17·55	1·78	19·7	—	—
Oberviechtach ...	15648	3293	3·94	18·69	1·24	25·8	—	—
Parsberg ...	28953	6825	4·56	19·36	2·45	40·1	—	—
Regensburg ...	29993	6665	4·36	19·63	3·98	37·6	81·8	967

* No particulars of Poverty Rate are available for the Bavarian Palatinate.

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

Appendix I, Principal Data etc. (continued).

Name of District	1 Population average, 1900—4	2 Number of women aged 16—50, Census 1901	3 Crude Birth Rate, 1900—4	4 Birth Rate in terms of Column 2	5 Poverty Rate, 1899—1902	6 Infant Death Rate, 1900—4	7 Percentage of children not Breast-fed, 1903—8	8 Number of children upon which (7) is based
Oberpfalz (*cont.*):								
Roding ...	23825	5194	4·14	18·99	3·93	30·4	54·3	313
Stadtamhof ...	42280	9690	5·01	21·87	3·64	40·4	56·9	810
Sulzbach ...	20338	4723	3·56	15·31	2·66	23·4	—	—
Tirschenreuth ...	35366	7717	3·47	15 90	2·41	18·1	—	—
Vohenstrauss ...	23787	5111	3·87	18·02	2·56	20·9	—	—
Waldmünchen ...	15831	3377	3·98	18·68	3·64	23 6	15·4	479
Oberfranken:								
Bamberg I ...	25347	5964	3·33	14·16	1·39	21·8	25·4	335
Bamberg II ...	28973	6433	3·56	16·04	1·55	23·0	11·5	286
Bayreuth ...	27700	6347	3 44	15 02	2·38	14·4	15·5	824
Berneck... ...	15271	3379	3·08	13·92	2·80	13·1	8·4	428
Ebermannstadt ...	22353	5253	3·03	12·89	1·86	15·8	9·4	595
Forchheim ...	28508	6655	3·18	13·63	1·60	17·1	—	—
Höchstadt a. A. ...	27516	6365	3·62	15·64	2·09	22·2	—	—
Hof	25962	5716	3·40	15·43	1·85	13·7	9·3	1389
Kronach ...	31684	7103	3·89	17·35	2·29	19·6	—	—
Kulmbach ...	25911	6003	3·11	13·44	1·72	13·2	11·3	785
Lichtenfals ...	33407	7901	3·55	15·01	1·58	21·2	—	—
Münchberg ...	27619	5982	3·07	14·17	2·08	12·0	4·8	769
Naila	22940	4910	3·22	15·04	1·32	13·2	10·0	619
Pegnitz	26456	6109	3·25	14·05	2·43	15·7	7·4	719
Rehau	25392	5442	3·26	15·19	1·62	14·8	—	—
Stadtsteinach ...	17252	3778	3·25	14·85	2·28	15·5	4·5	200
Staffelstein ...	18932	4569	3·02	12·51	1·82	20·6	22·6	336
Teuschnitz ...	18670	4150	3·87	17·42	0·86	19·9	5·2	419
Wunsiedel ...	45433	9757	3·40	15·81	1·97	13·7	11·6	595
Mittelfranken:								
Ansbach ...	33295	7894	3·09	13·02	1·72	20·4	18·6	307
Dinkelsbühl ...	24103	5726	3·28	13·81	2·14	28·3	—	—
Eichstätt ...	23752	5006	4·58	21·71	2·46	37·9	83·8	730
Erlangen ...	13221	3086	3·53	15·13	2·34	21·8	—	—
Feuchtwangen ...	26057	6064	3·36	14·45	3·05	26·0	23·0	665
Fürth	28397	6706	3·99	16·88	2·2[illegible]	26·6	50·6	160
Gunzenhausen ...	31938	7520	3·22	13·66	2·81	24·4	39·5	456
Hersbruck ...	41742	9283	3·32	14·94	2·53	21·5	—	—
Hilpoltstein ...	23550	5364	3·66	16·07	3·01	31·3	—	—
Neustadt a. A. ...	30044	6947	3·17	13·72	1·94	19·1	22·0	700
Nürnberg ...	20991	4703	3·93	17·55	1·89	26·7	45·6	193
Rothenburg a. T....	19623	4653	2·73	11·52	2·81	15·9	19·6	225
Scheinfeld ...	18935	4501	2·77	11·65	2·37	17·6	—	—
Schwabach ...	34270	8038	4·00	17·06	2·71	30·5	36·7	479
Uffenheim ...	30250	7267	2·55	10·67	2·62	17·6	37·8	579
Weissenburg i. B....	28012	6463	3·51	15·23	2·89	28·4	54·0	309

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

Name of District	1 Population average, 1900—4	2 Number of women aged 16—50, Census 1901	3 Crude Birth Rate, 1900—4	4 Birth Rate in terms of Column 2	5 Poverty Rate, 1899—1902	6 Infant Death Rate, 1900—4	7 Percentage of children not Breast-fed, 1903—8	8 Number of children upon which (7) is based
Unterfranken:								
Alzenau ...	22299	4908	3·84	17·45	0·89	14·0	—	—
Aschaffenburg ...	32071	7927	4·25	17·21	1·66	16·2	8·1	2048
Brückenau ...	12870	2840	3·27	14·80	1·38	13·7	21·1	346
Ebern	18900	4256	3·36	14·92	1·46	18·8	—	—
Gerolzhofen ...	30693	7112	3·16	13·64	2·81	19·0	16·6	313
Hammelburg ...	19631	4415	3·19	14·19	1·86	16·2	8·2	281
Hassfurt ...	28309	6479	3·82	16·69	2·07	21·9	29·5	536
Hofheim ...	13818	3310	3·00	12·52	0·88	18·0	24·9	329
Karlstadt ...	30199	7069	3·29	14·06	1·90	20·9	—	—
Kissingen ...	34473	8310	3·45	14·32	2·05	16·8	20·7	421
Kitzingen ...	29857	6985	2·78	11·89	2·68	17·7	17·7	265
Königshofen ...	14758	3498	3·11	13·12	1·56	19·3	—	—
†Lohr	19773 }	7995	3·74 }	15·62	2·08	20·5	17·3	1256
Gemünden ...	14853 }		3·42 }		1·95	16·2	16·2	401
Marktheidenfeld ...	30847	7061	3·37	14·71	1·83	17·9	—	—
Mellrichstadt ...	13426	3109	2·99	12·92	1·68	12·3	—	—
Miltenberg ...	21721	4832	3·37	15·16	1·59	16·7	—	—
Neustadt a. S. ...	19805	4620	3·35	14·38	2·69	17·4	—	—
Obernburg ...	27271	6045	3·35	15·09	1·57	14·4	3·7	540
Ochsenfurt ...	26625	6353	2·90	12·15	1·90	21·4	—	—
Schweinfurt ...	34703	8199	3·30	13·95	1·39	18·8	20·9	253
Würzburg ...	41728	9824	3·54	15·02	1·89	25·2	23·2	1008
Schwaben:								
Augsburg ...	40557	9507	4·17	17·80	2·22	34·3	68·6	1139
Dillingen ...	37507	8029	3·74	17·49	3·02	30·7	81·4	778
Donauwörth ...	32610	6582	3·76	18·61	2·23	33·1	—	—
Füssen	18805	4223	3·43	15·29	1·59	21·1	58·0	336
Günzburg ...	29905	6620	3·92	17·71	2·53	32·0	—	—
Illertissen ...	19606	4155	3·77	17·77	2·30	25·2	67·6	562
Kaufbeuren ...	23684	5142	3·61	16·61	1·57	26·1	—	—
Kempten ...	34397	7715	3·35	14·92	1·45	25·1	—	—
Krumbach ...	23849	5513	3·53	15·28	2·10	24·1	—	—
Lindau	31162	6936	3·05	13·71	1·40	18·6	—	—
Memmingen ...	30754	6761	3·46	15·76	1·33	23·2	71·1	422
Mindelheim ...	34094	7626	3·55	15·87	1·44	23·8	59·2	508
Neuburg a. D. ...	29319	6262	4·31	20·19	3·50	33·9	60·7	900
Neu-Ulm ...	20634	4562	3·59	16·25	2·05	24·8	46·3	227
Nördlingen ...	30675	7090	3·38	14·61	2·34	26·5	—	—
Oberdorf ...	24195	5233	3·19	14·74	1·14	22·1	65·4	680
Schwabmünchen ...	22305	4750	3·93	18·44	2·52	29·3	69·3	593
Sonthofen ...	34270	8235	2·96	12·33	1·32	18·9	57·7	307
Wertingen ...	18560	3936	3·91	18·47	2·55	29·1	65·9	1191
Zusmarshausen ...	16024	3437	3·95	18·42	2·13	27·2	65·8	433

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

APPENDIX II.

Table to show how the Means and Standard Deviations are affected by Variations in the numbers of Pairs available for different Correlations.

Character	Mean	Standard Deviation	Number of Observations
Infant Mortality Rate	24·309	7·642	161
,, ,, ,,	24·379	7·617	156
,, ,, ,,	25·183	7·468	145
,, ,, ,,	23·956 [24·917]	7·839 [7·592]	97
Birth Rate	16·429	2·589	156
,, ,, ...	16·481	2·574	142
,, ,, ...	16·225 [16·334]	2·616 [2·640]	92
Poverty Rate ...	2·408	·821	145
,, ,, ...	2·404	·821	142
,, ,, ...	2·427 [2·417]	·775 [·776]	87
Artificial Feeding Rate	43·367	28·355	97
,, ,, ,,	43·404	27·951	92
,, ,, ,,	46·685 [46·148]	27·999 [27·678]	87

[The values in brackets are those for 84 districts with records of each variable.]

https://doi.org/10.1017/S0022172400017009 Published online by Cambridge University Press

APPENDIX III.

The Effect on the Coefficient of Correlation of excluding a Proportion of the Frequency near the Mode of One Character.

Let the original correlation surface be normal, *i.e.*, let:

$$z = z_0 e^{-\frac{1}{2}(1-r^2)\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)}$$

where σ_1 and σ_2 are the standard deviations of the x and y characters and r the coefficient of correlation.

If z be multiplied by $\left(k - e^{-\frac{x^2}{2c^2}}\right)$ where k and c are constants, the former not less than unity, we shall be practising a method of exclusion which becomes more and more stringent as we approach the centroid. The coefficient of correlation deduced by product moments from this selected surface is given by the equation:

$$r_1 = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} kz\,xy\,dx\,dy - \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2c^2}} z\,xy\,dx\,dy}{\sqrt{\left\{\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} kx^2 z\,dx\,dy - \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2c^2}} x^2 . z\,dx\,dy\right)\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} ky^2 z\,dx\,dy - \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2c^2}} y^2 z\,dx\,dy\right)\right\}}}$$

we easily find:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2c^2}} xyz\,.\,dx\,dy = \frac{N\sigma_1\sigma_2 rc^3}{(\sigma_1^2+c^2)^{\frac{3}{2}}}$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2c^2}} x^2 z\,dx\,dy = \frac{N\sigma_1^2 c^3}{(\sigma_1^2+c^2)^{\frac{3}{2}}}$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2c^2}} y^2 z\,dx\,dy = \frac{N\sigma_2^2 c\,(\sigma_1^2+c^2-r^2\sigma_1^2)}{(\sigma_1^2+c^2)^{\frac{3}{2}}}.$$

The other three integrals are, of course, $kN\sigma_1\sigma_2 r$, $kN\sigma_1^2$ and $kN\sigma_2^2$ respectively and we find on substituting:

$$r_1 = r \times \sqrt{\frac{k(\sigma_1^2+c^2)^{\frac{3}{2}} - c^3}{k(\sigma_1^2+c^2)^{\frac{3}{2}} - c^3 - \sigma_1^2 c\,(1-r^2)}} \quad \text{..........................(1)}$$

i.e., r_1 is numerically greater than r for all values of the latter between 0 and $\pm$ 1.

My friend Mr G. Udny Yule, with whom I have discussed the problem referred to in this note, has pointed out to me that a result identical with (1) can be deduced from Pearson's (1903) constants for a selected population obeying the normal law. The distribution here considered will be the original population of Pearson's case less the normal selection (*e.g.*, if the selected individuals represent survivors the distribution will be that of deaths). Thus, Mr Yule finds:

"SELECTION." "REST."

$$N' = N \times \frac{yc}{\sqrt{c^2+\sigma_1^2}} \qquad N'' = N\left(1 - \frac{yc}{\sqrt{c^2+\sigma_1^2}}\right)$$

$$S_1 = \sigma_1 \times \frac{c}{\sqrt{c^2+\sigma_1^2}} \qquad S_1'^2 = \frac{\sigma_1^2}{c^2+\sigma_1^2} \times \frac{(c^2+\sigma_1^2)^{\frac{3}{2}} - yc^3}{\sqrt{(c^2+\sigma_1^2)} - yc} \text{ [since } N'S_1^2 + N''S_1'^2 = N\sigma_1^2\text{]}$$

$$S_2 = \sigma_2 \times \frac{\sqrt{c^2 + (1-r^2)\sigma_1^2}}{\sqrt{c^2+\sigma_1^2}} \qquad S_2'^2 = \frac{\sigma_2^2}{c^2+\sigma_1^2} \times \frac{(c^2+\sigma_1^2)^{\frac{3}{2}} - yc^3 - yc\sigma_1^2(1-r^2)}{\sqrt{(c^2+\sigma_1^2)} - yc}$$

$$p' = R's_1s_2 = p \times \frac{c^2}{c^2+\sigma_1^2} \qquad p'' = \frac{p}{c^2+\sigma_1^2} \times \frac{(c^2+\sigma_1^2)^{\frac{3}{2}} - yc^3}{\sqrt{(c^2+\sigma_1^2)} - yc}$$

$$R' = r \times \frac{c}{\sqrt{c^2 + (1-r^2)\sigma_1^2}} \qquad R'' = r \times \sqrt{\frac{(c^2+\sigma_1^2)^{\frac{3}{2}} - yc^3}{(c^2+\sigma_1^2)^{\frac{3}{2}} - yc^3 - yc\sigma_1^2(1-r^2)}}$$

where σ, s and S' represent the standard deviations, N, N' and N'' the populations, r, R' and R'' the coefficients of correlation, p, p' and p'' the product moments of the whole population, "Selection" and "Rest" respectively, while y is a constant less than unity.

It will be seen that Mr Yule's R'' agrees with (1) *supra* when we substitute $\frac{1}{y}$ for k in the latter.

The point is a simple one, but I do not remember to have seen it dealt with elsewhere.