Chapter IV

MALTHUSIAN POPULATIONS WITH KNOWN AGE DISTRIBUTION: SUB-SETS F(r)

We propose to study in this chapter the sub-sets F(r) which were defined in chapter I as a kind of counterpart of the sub-sets H(r). If we assume that the age distribution $C_0(a)$ is *constant and given*, while the survivorship function $p(a)$ is constant but not known, then all the Malthusian populations satisfying these conditions form the sub-set $F_0(r)$ corresponding to the age distribution $C_0(a)$, and by varying $C_0(a)$ we obtain the series of sub-sets F(r).

A. The fundamental formulae

The basic formulae for the sub-sets F(r) are the same as for the sub-sets H(r). We shall set them out once again, using a particular sub-set $F_0(r)$, in a form expressing the unknown functions in terms of the age distribution $C_0(a)$.

(1) Let us note, first of all, that $C_0(0) = b_0$. Consequently, when we assume a given age distribution we also assume a corresponding crude birth rate;

(2) The survivorship function is written:

$$p(a) = \frac{C_0(a)}{C_0(0)} e^{ra} \qquad \text{(II.3}d\text{)}^1$$

and thus we obtain all the survivorship functions of the sub-set $F_0(r)$ by successively giving r all its possible values;

(3) The rate r must be such that the function $p(a)$ thus calculated is a decreasing function. We therefore must have:

$$\frac{dp(a)}{da} < 0$$

or, in other terms,

$$r < -\frac{C_0'(a)}{C_0(a)}$$

Let

$$-\frac{C'(a)}{C(a)} = Q(a)$$

r must therefore satisfy the inequality:

$$r < Q_0(a) \qquad \text{(IV.1)}$$

For a given function $C_0(a)$, $Q_0(a)$ passes through a minimum for a certain age a_m. This minimum thus represents a maximum for r. Finally, we can arrive at all the populations of the sub-set F(r) by varying r from $-\infty$ to $Q_0(a_m)$.[2]

(4) The probabilities of death given in the various model life tables calculated with the aid of formula II.3d have an important property, since we have:

$$q(a) = Q_0(a) - r \qquad \text{(II.4)}$$

Each series of probabilities is obtained by subtracting r from the series of probabilities $Q_0(a)$.

(5) The fertility function $\varphi(a, t)$ satisfies the condition:

$$\int_0^\infty \varphi(a,t) C_0(a)\,da = b_0 = C_0(0) \qquad \text{(II.15)}$$

It may be recalled that in a sub-set H(r) we could arbitrarily assume the function $\varphi(a, t)$. This value then determined a particular Malthusian population which we called a stable population. This no longer applies in the case of a sub-set F(r), however. Here, we cannot arbitrarily assume the fertility function, which must satisfy II.15.

B. Populations of a sub-set F(r) satisfying certain conditions

In order to define a particular population of the sub-set F(r), we must have an additional condition. We shall now discuss some of these conditions, which are of particular interest in practice.

FIRST EXAMPLE: MALTHUSIAN POPULATION WITH AGE DISTRIBUTION $C_0(a)$ AND GIVEN RATE OF VARIATION r_0

The simplest way to define a particular population of the sub-set F(r) is obviously to assume the rate of variation r_0 as given. We then apply formula II.3d to calculate the corresponding survivorship function.

$$p_0(a) = \frac{C_0(a)}{C_0(0)} e^{r_0 a} \qquad \text{(II.3}d\ bis\text{)}$$

The numerical calculations raise various practical problems which we shall now examine.

Discontinuous data. The age distribution $C_0(a)$ is generally not known in continuous form but is, rather, a function which is known for various age groups. We shall therefore take the following conventional age groups : 0 (i.e., less than 1 year of age),[3] 1-4, 5-9, 10-14, and so on by successive five-year age groups. The last age group usually consists of several five-year groups together. For clarity, we shall make this last age group cover all persons aged 85 and over.

[1] This is formula II.3d in table II.2.

[2] This is obviously only a theoretical possibility. We have already pointed out that in actual populations other limits might apply.

[3] We shall consider later the case where the 0 and 1-4 age groups are combined into a single age group 0-4.

One of the first consequences of the fact that we know the age distribution in discontinuous notation is the following: in continuous notation, we had $C_0(0) = b_0$, thus knowledge of the age distribution automatically meant that we also knew the crude birth rate; in discontinuous notation, however, with the age groups considered here, this no longer applies and the crude birth rate is unknown.

In order to apply formula II.3*d bis*, we use approximate relationships based on the following assumptions:

(*a*) For a five-year age group a, $a + 5$ we assume that:[4]

$$b_0 L_a = C_a e^{r(a+2.5)}$$

For the first two age groups, we assume that:

$$b_0 L_0 = C_0 e^{0.5r}$$

and

$$b_0 L_{1\text{-}4} = C_{1\text{-}4} e^{3r}$$

Finally, for the last age group, we assume that:

$$b_0 L_{85+} = C_{85+} e^{87.5r}$$

These formulae make it possible to calculate the series:

$$b_0 L_0,\ b_0 L_{1\text{-}4},\ b_0 L_{5\text{-}9},\ \ldots\ b_0 L_{85+}$$

(*b*) On the basis of this series, we can calculate the series of values of $b_0\ p_0(a)$ by applying the following formulae:

(1) For each five-year age group—e.g., L_{10-14}—we assume that:

$$L_{10\text{-}14} = \frac{5}{2}[p(10) + p(15)]$$

We also have the approximate formula:

$$p(10) + p(15) = 2p(12.5)$$

so that:

$$L_{10\text{-}14} = 5p(12.5)$$

Similarly, we can write:

$$L_{5\text{-}9} = 5p(7.5)$$

and finally:

$$L_{10\text{-}14} + L_{5\text{-}9} = 5p(12.5) + 5p(7.5) = 10p(10)$$

We can thus very easily calculate the series:

$$b_0 p(7.5),\ b_0 p(10),\ b_0 p(12.5),\ \ldots\ b_0 p(82.5)$$

Finally, the formula:

$$L_{5\text{-}9} = \frac{5}{2}[p(5) + p(10)]$$

enables $b_0 p(5)$ to be calculated.

(2) For the last age group, we apply the formula:[5]

$$L_{85+} = p(85) \text{ Log } 100{,}000 p(85)$$

We determine $b_0 p(85)$ by a process of trial and error.

(3) It now remains for us to find the values of the survivorship function under the age of 5. For the 1-4 age group, we write:

$$L_{1\text{-}4} = 1.9p(1) + 2.1p(5)$$

As we already know $b_0 p(5)$, we can calculate $b_0 p(1)$ with the aid of this formula.

The problem of the 0 age group. In the case of the 0 age group the problem is more complicated; for, if λ represents the proportion of deaths of children under 1 year of age which occur in the calendar year of their birth, we have the following approximate formula:

$$L_0 = 1 - \lambda + \lambda p(1)$$

which can be written:

$$b_0 L_0 = C_0 e^{0.5r} = b_0(1 - \lambda) + b_0 \lambda p(1)$$

from which we have:

$$b_0 = \frac{C_0 e^{0.5r} - \lambda b_0 p(1)]}{1 - \lambda} \qquad \text{(IV.2)}$$

We already know $b_0 p(1)$, but we do not know λ to complete the calculation. We again come up against the fact mentioned above, namely, that knowledge of the age distribution in discontinuous terms does not mean that we know the crude birth rate.

The proportion λ is not the same for all populations. In practice[6] it varies from 0.6 to 0.9. By varying λ between these limits, we can therefore determine with the use of formula IV.2 a whole series of crude birth rates.

As long as r remains small—as it always does in the case of the human species—the results scarcely depend on the value of r adopted for the computation. We can therefore speak of crude birth rates compatible with the age distribution, and it is useful to determine this series of compatible crude birth rates. If b_0 is given as well as the age distribution, we can check that it is indeed within the series; when it is not given, we merely have to select of value of b_0 which is within the series in order to complete the computation.

Use of the infant mortality. Determination of the compatible crude birth rates is tantamount to choosing the infant mortality measured for the probability of death between the ages of 0 and 1. In fact, choosing b_0 when $b_0 p(1)$ is known is tantamount to choosing the infant probability of death q_0.

---

[4] It may be recalled that:

$$L_a = \int_a^{a+5} p(a)da$$

[5] This formula is given in the manual on the computation of population projections, *Manual III. Methods for Population Projections by Sex and Age* (United Nations publication, Sales No.: 56.XIII.3). It is an experimental formula which is found to be satisfied within the range of variations of mortality of the human species.

[6] It is convenient to draw a distinction between the two main categories of infant deaths: endogenous deaths, which are due to factors present when the infant is born (genetic factors arising at conception, factors transmitted by the mother during pregnancy, or factors arising at confinement), and exogenous deaths, which are due to the environment—in the broad sense of the term—in which the infant lives after birth. Endogenous deaths occur for the most part soon after birth, and they can be considered in practice to occur entirely within the calendar year of birth, but exogenous deaths occur throughout the calendar year of birth and the following year. Observation reveals that this distribution is remarkably stable in time and space. If e is the endogenous infant mortality (per 1,000 births) and E is the exogenous infant mortality, distributed between kE, the calendar year of birth, and $(1 - k)$E, the following year, we can write:

$$\lambda = \frac{e + kE}{e + E}$$

where k is a coefficient which is the same for all populations. λ is then a function of e and E. On the basis of observed values of e and E, we find λ to be between 0.6 and 0.9.

As we are assuming that the mortality is that of the human species, we know that in practice this probability can only vary between 10 and 500 per thousand and that the value of q_0 selected must come within these limits. By using the formulae given above, we can easily see that we have:

$$\lambda = \frac{b_0 p(1) - C_0 e^{0.5}(1 - q_0)]}{q_0 b_0 p(1)} \qquad \text{(IV.3)}$$

Once we have chosen q_0, therefore, we can calculate λ, which must be between 0.6 and 0.9. If λ is outside these limits, the value originally chosen for q_0 must be rejected. We can thus determine a range of variation for infant mortality compatible with the age distribution, just as we did for the crude birth rate. The range thus determined is generally smaller than the range of 10 to 500 per thousand which was quoted above as representing the practical limits of infant mortality in the human species. We can, in fact, write formula (IV.3) as follows:

$$q_0 = \frac{1 - \dfrac{C_0 e^{0.5r}}{b_0 p(1)}}{\lambda - \dfrac{C_0 e^{0.5r}}{b_0 p(1)}}$$

In this form, we can easily calculate the limits of variation of q_0 by successively taking $\lambda = 0.6$ and $\lambda = 0.9$.

When the first age group is 0-4 years. It may happen that the age distribution is not known separately for age groups 0 and 1-4, but only for the combined 0-4 age group. As in the previous case, we first of all determine the series:

$$b_0 L_{0\text{-}4},\ b_0 L_{5\text{-}9},\ b_0 L_{10\text{-}14} \text{ etc.}$$

by the formulae:

$$C_{0\text{-}4} e^{2.5r},\ C_{5\text{-}9} e^{7.5r},\ C_{10\text{-}14} e^{12.5r} \text{ etc}$$

From this, we have no difficulty in progressing, as described above, to the series: $b_0 p(5)$, $b_0 p(10)$, $b_0 p(15)$ etc. and in order to complete the calculation we must, as before, assume a value for b_0, or for the infant mortality q_0. Can we also calculate for these two quantities what we called above a "range compatible with the age distribution"? We can write: $b_0 p(1) = b_0(1 - q_0)$ and in accordance with the foregoing formulae:

$$b_0 L_0 = b_0(1 - \lambda) + \lambda b_0 p(1) = b_0(1 - \lambda) + \lambda b_0(1 - q_0)$$

$$b_0 L_{1\text{-}4} = 1.9 b_0 p(1) + 2.1 b_0 p(5) = 1.9 b_0(1 - q_0) + 2.1 b_0 p(5)$$

Whence we finally have:

$$b_0 L_{0\text{-}4} = b_0 L_0 + b_0 L_{1\text{-}4} = b_0(1 - \lambda) + \lambda b_0(1 - q_0) + 1.9 b_0(1 - q_0) + 2.1 b_0 p(5)$$

We know that: $b_0 L_{0\text{-}4} = C_{0\text{-}4} e^{2.5r}$

We therefore have the equation:

$$C_{0\text{-}4} e^{2.5r} = b_0(1 - \lambda) + \lambda b_0(1 - q_0) + 1.9 b_0(1 - q_0) + 2.1 b_0 p(5) \qquad \text{(IV.4)}$$

Let us suppose that q_0 is given. We then have a relationship between b_0 and λ, in which to each value for λ corresponds one value for b_0. Experience shows that, if λ varies from 0.6 to 0.9, the corresponding values for b_0 will not be very different. If q_0 varies from 10 to 500 per thousand and λ from 0.6 to 0.9, we obtain a series of rates b_0 which vary within two limits. These limits define the range of variation of b_0 compatible with the age distribution. The result is practically independent of the value of r used in the calculations.

In the preceding case, where we knew the age distribution of the 0-1 and 1-4 age groups, we completed the calculation by assuming a value for the infant mortality q_0, and λ was then determined. This determination of λ amounted, in fact, to the determination of the distribution of q_0 between endogenous and exogenous causes of death.

In the present case, where we know only the age distribution for the 0-4 age group, assumption of a value for the infant mortality q_0 will only enable us to complete the calculation with a certain margin of error which is fortunately quite small, corresponding to the possibility of variation of λ between 0.6 and 0.9. In order to eliminate this possible error, we must also assume a value for λ, i.e., we must select a value for the distribution of the infant mortality between endogenous and exogenous causes of death.

In the first case knowledge of λ was enough to rule out any question of a choice, but in the second case knowledge of λ does not eliminate the need to choose a value for q_0 and only removes a secondary uncertainty. These explanations will be made clearer and more specific if we give some numerical applications.

A numerical application. In this first example, we shall use the age distribution of the female Malthusian population having a rate of natural variation $r = 0.03$ and a mortality identical with that of the intermediate model life table giving an expectation of life at birth for both sexes of 50 years. This age distribution has been computed in table II.3 and is given again in table IV.2.

Leaving aside the way in which this age distribution has been computed and taking it as a known quantity, we shall consider the sub-set $F_0(r)$ which corresponds to it.

We shall first of all determine the series of crude birth rates compatible with this age distribution, by applying formula IV.2. As already stated, the result does not in practice depend on the value adopted for r. In order to verify this fact, we made two calculations in which we successively took the value of r as 0 and 0.015. Table IV.1 shows that we do indeed obtain two almost identical series of compatible rates. In practice, of course, we should only made a single calculation, and the simplest procedure would be to assume that $r = 0$.

Let us now examine the computation for this case in detail:[7]

$$b_0 p(10) = C(10) = \frac{C_{5\text{-}9} + C_{10\text{-}14}}{10} = \frac{146{,}358 + 123{,}799}{10} = 27{,}015.7$$

$$b_0 p(7.5) = C(7.5) = \frac{C_{5\text{-}9}}{5} = \frac{146{,}358}{5} = 29{,}271.6$$

As C(5) is symmetrical with C(10) in relation to C(7.5), we have:

$$(C\ 5) = 29{,}271.6 + (29{,}271.6 - 27{,}015.7) = 31{,}527.5$$

[7] The numerical values used in this computation are given in table IV.2.

Let us now proceed to the computation of $b_0C(1)$. We have:

$$C_{1\text{-}4} = b_0L_{1-4} = 1.9b_0p(1) + 2.1b_0p(5) = 139{,}640$$

whence we have:

$$b_0p\,(1) = 38{,}648.6$$

We can then apply formula IV.2 for various values of r. The results are given in the column headed $r = 0$ in table IV.1. The same calculation for $r = 0.015$ gives the series of crude birth rates appearing in the column headed $r = 0.015$. As already stated, the two series are almost identical.

TABLE IV.1. CRUDE BIRTH RATE PER 1,000 COMPATIBLE WITH THE AGE DISTRIBUTION OF THE POPULATION IN TABLE IV.2 (COLUMN 3)

a	$r = 0$	$r = 0.015$
0.6	42.7	42.9
0.7	44.1	44.1
0.8	46.8	46.6
0.9	55.0	54.2

We now propose to find the population of sub-set $F_0(r)$ corresponding to $r = 0.015$. Table IV.2 gives the details of the computation.

Column 1 of the table gives the median age a of the age groups whose limits are indicated in column 2. The age distribution of the population, C_a, which is the same as the age distribution of the population computed in table II.3, is given in column 3. Column 4 gives the sequence of the coefficients $e^{r\alpha}$ for $r = 0.015$. These are the data of the problem. The computations are shown from column 5 onwards.

We begin by calculating the sequence of the quantities b_0L_a. These are given in column 5 and are obtained by multiplying columns 3 and 4.

Column 6 gives the sequence of the quantities b_0, $b_0p(1)$, $b_0p(5)$, $b_0p(10)$, $b_0p(15)$, and so forth. These are obtained in the following manner:

(*a*) The figures below the horizontal line are obtained by successively adding together in pairs the figures in column 5 and dividing the result by 10. Thus, the first figure below the line comes to:

$$\frac{(163{,}799 + 149{,}330)}{10} = 31{,}312.9 = b_0p(10)$$

b) For $b_0p(5)$, we have:

$$b_0p(5) = 32{,}759.3 + (32{,}759.3 - 31{,}312.9) = 34{,}206.8$$

(*c*) For $b_0p(1)$, we have:

$$1.9b_0p(1) + 2.1b_0p(5) = 146{,}063$$

TABLE IV.2. COMPUTATION OF THE SURVIVORSHIP FUNCTIONS CORRESPONDING IN A MALTHUSIAN POPULATION TO A GIVEN AGE DISTRIBUTION C_a AND AN INTRINSIC RATE OF NATURAL VARIATION OF 0.015

Median age a (1)	*Age group (years)* (2)	*Age distribution* [a] C_a (3)	e^{ra} *for* $r = 0.015$ (4)	*Product of the two preceding columns:* $C_ae^{ra} = b_0L_a$ (5)	$\frac{b_0L_a + b_0L_{a+5}}{10} = b_0p(a+5)$ (6)	*Differences between successive figures in the preceding column* (7)	*Death rate* m_a *(per thousand); column (7) divided by column (5)* (8)	*Survivors to the initial age of the age group; column (6) divided by 45,074.6 and multiplied by 100,000* (9)
0.5	0	40 282	1.00750	40 584	45 074.6	6 006.8	147.53	100 000
3.0	1-4	139 640	1.04600	146 063	39 067.7	4 860.9	33.53	86 717
7.5	5-9	146 358	1.11917	163 799	34 206.8	2 893.9	17.56	75 851
12.5	10-14	123 799	1.20623	149 330	31 312.9	2 765.5	18.52	69 469
17.5	15-19	104 623	1.30128	136 144	28 547.4	2 634.8	19.36	63 334
22.5	20-24	87 754	1.40144	122 982	25 912.6	2 553.6	20.76	57 488
27.5	25-29	73 222	1.51059	110 608	23 359.0	2 370.8	21.43	51 823
32.5	30-34	60 968	1.62829	99 274	20 988.2	2 169.0	21.86	46 563
37.5	35-39	50 661	1.75515	88 918	18 819.2	1 990.2	22.39	41 751
42.5	40-44	41 957	1.89174	79 372	16 829.0	1 854.8	23.37	37 336
47.5	45-49	34 509	2.03918	70 370	14 974.2	1 773.7	25.19	33 221
52.5	50-54	28 029	2.19899	61 635	13 200.5	1 749.0	28.38	29 286
57.5	55-59	22 320	2.36918	52 880	11 451.5	1 772.9	33.52	25 406
62.5	60-64	17 193	2.55369	43 906	9 678.6	1 842.3	41.98	21 472
67.5	65-69	12 518	2.75257	34 457	7 836.3	1 926.0	55.90	17 385
72.5	70-74	8 307	2.96685	24 646	5 910.3	1 922.5	78.00	13 112
77.5	75-79	4 763	3.19792	15 232	3 987.8	1 705.9	111.97	8 847
82.5	80-84	2 201	3.44709	7 587	2 281.9	1 287.1	169.68	5 063
87.5	85-89	896	3.71655	3 330	994.8	994.8	298.97	2 207
TOTAL . . .				1 451 088				

[a] Figures taken from table II.3 (chap. II).

As we already know $b_0p(5)$, we obtain from this equation: $b_0p(1) = 39{,}067.7$.

(*d*) Finally, in order to complete the computation, we must select a value of b_0 from the series of compatible crude birth rates. We have selected $b_0 = 0.0450746$, which is the crude birth rate of the Malthusian population constructed as a first step in table II.3 in order to obtain the age distribution C_a. We need not have made this particular choice, but could have adopted any value compatible with the age distribution.

Column 7 gives the differences between successive figures in column 6.

Column 8 gives the death rates by age groups m_a. These are obtained by dividing column 7 by column 5.

Finally, column 9 gives the survivors to the initial ages of the age groups: $p(0)$, $p(1)$, $p(5)$, $p(10)$, $p(15)$, and so forth.

It should be noted that it is not necessary to know b_0 in order to calculate the survivorship function from 1 year of age onwards. All that is necessary is to divide the terms of the series $b_0p(1)$, $b_0p(5)$, $b_0p(10)$, $b_0p(15)$, and so forth, by $b_0p(1)$.

Similarly, all the death rates m_a except the first can be calculated without reference to b_0.

Let us now suppose that we know only the 0-4 age group. Table IV.3 is then written as follows:

TABLE IV.3

Median age a (1)	Age group (years) (2)	Age distribution of the population, C_a (3)	e^{ra} (4)	Product of the two proceding columns, $C_a e^{ra} = b_0L_a$ (5)	$\frac{b_0L_a + b_0L_{a+5}}{10} = b_0p(a+5)$ (6)
2.5	0-4	179 922	1.03821	186 797	
7.5	5-9	146 358	1.11397	163 799	34 206.8
12.5	10-14	123 799	1.20623	149 330	31 312.9
17.5	15-19	104 623	1.30128	136 144	28 547.4
. . . .	...	...	...	...	...
. . . .	...	...	...	...	...
. . . .	...	...	...	...	...

Over 4 years of age, nothing needs to be changed. Only the first line, which is a combination of the first two lines of table IV.2, is changed, so that we have:

$$C_{0\text{-}4}e^{2.5r} = 186{,}797$$

and $2.1b_0p(5) = 2.1 \times 34{,}206.8 = 71{,}834.3$.

Formula IV.4 is written:

$$\frac{(186{,}797 - 71{,}834)}{1-\lambda+\lambda(1-q_0)+1.9(1-q_0)} = b_0$$

or

$$b_0 = \frac{114{,}963}{2.9-(\lambda+1.9)q_0}$$

This formula enables b_0 to be computed, once λ and q_0 are known. Table IV.4 gives the results of this computation. It was stated above that the result depends scarcely at all on the value of r used. This can easily be verified in the following table, which shows two series of compatible crude birth rates corresponding to $r = 0$ and $r = 0.015$. The two series are almost identical.

TABLE IV.4. CRUDE BIRTH RATES COMPATIBLE WITH THE AGE DISTRIBUTION IN TABLE IV.2, FOR VARIOUS LEVELS OF INFANT MORTALITY, TWO VALUES OF λ AND TWO RATES OF NATURAL VARIATION (APPLICATION OF FORMULA IV.4)

q_0 (rate per thousand)	$\lambda = 0.6$		$\lambda = 0.9$	
	$r = 0.015$	$r = 0$	$r = 0.015$	$r = 0$
10	40.0	39.6	40.0	39.6
50	41.4	41.0	41.7	41.3
100	43.4	43.0	43.9	43.4
200	47.9	47.4	49.1	48.6
300	53.5	52.9	55.8	55.2
400	60.5	59.8	64.6	63.9
500	69.7	68.9	76.6	75.8

It can be seen very clearly that the variation of λ has little effect on b_0. Finally, the crude birth rates compatible with the age distribution vary from 40.0 to 76.6. We may recall that for the same age distribution (but on the assumption that we knew both the 0 and the 1-4 age groups) we found compatible crude birth rates varying from 42.7 to 55.0.

Direct computation of probabilities of death over a finite age range. We can obviously determine the probabilities of death from the survivorship table in table IV.2. These probabilities can also be obtained directly, however, without making use of the survivorship table.

If we divide each five-year age group by the group which precedes it, we can write:

$$e^{5r}\frac{C_{a+5}}{C_a} = \frac{L_{a+5}}{L_a} = \frac{p(a+7.5)}{p(a+2.5)}$$

We thus obtain the survivorship rates for:

from 7.5 to 12.5 years of age
from 12.5 to 17.5 years of age
from 17.5 to 22.5 years of age
... ...
from 77.5 to 82.5 years of age
from 82.5 to 85 and over

For earlier ages, we can write:

$$e^{4.5r}\frac{C_{5\text{-}9}}{C_{1\text{-}4}}\frac{4}{5} = \frac{p(7.5)}{p(3)}$$

i.e., the survivorship rate from 3 to 7.5 years of age, and likewise:

$$e^r \frac{C_{1-4}}{C_0} \frac{1}{4} = \frac{p(3)}{p(0.5)}$$

i.e., the survivorship rate from 0.5 to 3 years of age, and finally:

$$e^r \frac{C_0}{b_0} = \frac{p(0.5)}{p(0)}$$

i.e., the survivorship rate from birth to 0.5 years of age.

The complements to 1 of all these survivorship rates are the probabilities of death calculated for finite age ranges: 0.5 years in the case of the first age group, 2.5 years for the second, 4.5 years for the third and 5 years for all the others. All this clearly shows that we must know b_0 in order to obtain the complete life table: otherwise, it is possible to determine the probabilities only from the age of 0.5 years.

An important property of the survival ratios. We see that the survival ratios for five-year periods have a remarkable property: over the age of 7.5 years, they can be deduced from the series of coefficients

$$\frac{C_a + 5}{C_a}$$

by multiplying by e^{5r}.

This remarkable property corresponds to that already pointed out above in the case of the instantaneous probabilities. It will be recalled that these probabilities were deduced from Q(a) (the logarithmic derivative bearing the sign of C(a)) by taking away a constant, r.

$$q(a) = Q(a) - r$$

In the case of the finite probabilities of death, this property is only approximate. Thus, we have:

$$_5q_a = \frac{p(a) - p(a+5)}{p(a)} =$$

$$= \frac{C(a) - C(a+5)e^{5r}}{C(a)} = 1 - \frac{C(a+5)}{C(a)} e^{5r}$$

We also have:

$$_5Q_a = \frac{C(a) - C(a+5)}{C(a)} = 1 - \frac{C(a+5)}{C(a)}$$

whence we have the difference:

$$_5Q_a - {_5q_a} = \frac{C(a+5)}{C(a)} (e^{5r} - 1)$$

For a given value of r, this difference varies with age, since $C(a+5)/C(a)$ is not constant. If we set aside the extreme ages of life, however, $C(a+5)/C(a)$ remains in reality close to unity, and as r is relatively small, we have approximately:

$$_5q_a \# {_5Q_a} - 5r$$

i.e., an approximate formula similar to the exact formula regarding the instantaneous probabilities.

The condition of compatibility of r. As was seen above, not all values of r are compatible with a given age distribution. r must therefore satisfy the condition:[8]

$$r < -\frac{C'(a)}{C(a)}$$

This condition is easily expressed in discontinuous notation when considering the survivorship rates. It is tantamount to saying that survivorship rates computed in the manner described above must all be less than unity.

Numerical example of the direct computation of survival ratios. Let us again take the age distribution in table II.3 and compute the life table corresponding to $r = 0.03$. It will be noted that this is exactly the same value as was used for the original construction of the age distribution in table II.3. If, therefore, we take $b_0 = 0.0450751$, we should obtain the same life table as was used with $r = 0.03$ for that construction. In fact, however, as we are using approximate formulae, we cannot hope to obtain exactly the same life table as in the first case, and indeed the differences enable us to evaluate the usefulness of the approximate formulae.

Table IV.5 gives the details of the computation, which does not call for any special comments, since it is the direct application of the formulae already given. A comparison of columns 10 and 11 and of columns 12 and 14 shows that the approximate formulae give very good results.

Determination of the values of r *compatible with the age distribution.* We shall, however, describe at this point how the computations in table IV.5 enable the values of r compatible with the age distribution to be determined. We have already seen that the condition of compatibility is tantamount to saying that the survivorship rates must all be less than unity. These survivorship rates, for $r = 0.03$, are given in column 6. They pass through a maximum equal to $0.84586 \times e^{5 \times 0.03}$ for the 5-9 age group, whatever the value of r. We should therefore have: $0.84586e^{5r} < 1$, which gives: $r < 0.033480$.

Thus, all the values of r below this limit of $r_m = 0.033480$ are compatible with the age distribution computed in table IV.5.

Verification of the important property of the instantaneous death rates. If we leave aside the extreme ages, the death rate for an age group m_a is close to the probability of death at the median age of the age group. Consequently, the two series of death rates in tables IV.2 and IV.5 should illustrate in an approximate manner the remarkable property of the instantaneous death rates mentioned above. There should be a difference between the two series, at all ages, of about 0.015,[9] and this is in fact exactly what we observe in table IV.6.

Numerical application on the basis of an actually observed age distribution. We now propose to repeat the same calculations, by using the age distributions of actual populations. We shall begin by using an age distribution

[8] In the example in table III.2, we did not take the trouble to verify that this condition was fulfilled because we were quite sure that it was, since the age distribution used had been constructed as a first step, by using the value $r = 0.03$. When using this age distribution in the opposite sense, we were sure that at least all the values of r equal to or less than 0.03 would be compatible with this age distribution.

[9] 0.015 is the difference between the rate of $r = 0.015$ used in table IV.2 and the rate of $r = 0.030$ used in table IV.5.

TABLE IV.5. COMPUTATION OF THE MORTALITY RATES CORRESPONDING, IN A MALTHUSIAN POPULATION, TO A GIVEN AGE DISTRIBUTION C_a AND AN INTRINSIC RATE OF NATURAL VARIATION OF $r = 0.03$ (DIRECT CALCULATION OF THE PROBABILITIES OF DEATH)

Median age a *(1)*	*Age group (years) "a"* *(2)*	*Age distribution of the population* C_a *(3)*	$\frac{C_a + 5}{C_a}$ *(4)*	e^{ra} *for* $r = 0.03$ (c) *(5)*	*Survivorship rate* $\frac{L_a + 5}{L_a}$ *(6)*	*Probability of death between one age and the following age (per 1,000)* *(7)*	*Survivors to age "a"* $p(a)$ *(8)*	*Deaths between age "a" and the following age* *(9)*	*Survivors in each age group*: L_a *reconstitued* *(10)*	*Survivors in each age group*: L_a *initial* *(11)*	*Death rate (per 1 000)*: m_a *reconstitued* *(12)*	*Death rate (per 1 000)*: m_a *initial* *(13)*
0 . . .	Births	45 075 (a)	0.89366 (b)	1.0151 (c)	0.90715	92.85	100 000	9 285				
0.5 . . .	Under 1 year	40 282	0.86664	1.0779	0.93415	65.85	90 715	5 974	90 715	90 719	136.41	136.41
3.0 . . .	1-4	139 640	0.83849	1.1445	0.95965	40.35	84 741	3 419	338 964	338 974	16.19	16.19
7.5 . . .	5-9	146 358	0.84586	1.1618	0.98272	17.28	81 322	1 405	406 610	406 628	3.72	3.99
12.5 . . .	10-14	123 799	0.84510	1.1618	0.98184	18.16	79 917	1 451	399 585	399 620	3.57	2.96
17.5 . . .	15-19	104 623	0.83876	1.1618	0.97447	25.53	78 466	2 003	392 330	392 370	4.40	4.38
22.5 . . .	20-24	87 754	0.83440	1.1618	0.96941	30.59	76 463	2 339	382 315	382 368	5.67	5.97
27.5 . . .	25-29	73 222	0.83265	1.1618	0.96737	32.63	74 124	2 419	370 620	370 680	6.41	6.45
32.5 . . .	30-34	60 968	0.83094	1.1618	0.96539	34.61	71 705	2 482	358 525	358 600	6.83	6.80
37.5 . . .	35-39	50 661	0.82819	1.1618	0.96219	37.81	69 223	2 617	346 115	346 202	7.36	7.28
42.5 . . .	40-44	41 957	0.82248	1.1618	0.95556	44.44	66 606	2 960	333 030	333 118	8.37	8.15
47.5 . . .	45-49	34 509	0.81222	1.1618	0.94364	56.36	63 646	3 587	318 230	318 325	10.28	10.06
52.5 . . .	50-54	28 029	0.79632	1.1618	0.92516	74.84	60 059	4 495	300 295	300 392	13.45	13.22
57.5 . . .	55-59	22 320	0.77030	1.1618	0.89493	105.07	55 564	5 838	277 820	277 922	18.59	18.05
62.5 . . .	60-64	17 193	0.72809	1.1618	0.84589	154.11	49 726	7 663	248 630	248 722	27.15	26.79
67.5 . . .	65-69	12 518	0.66360	1.1618	0.77097	229.03	42 063	9 634	210 315	210 400	41.12	41.19
72.5 . . .	70-74	8 307	0.57337	1.1618	0.66614	333.86	32 429	10 827	162 145	162 220	63.08	65.38
77.5 . . .	75-79	4 763	0.46210	1.1618	0.53687	463.13	21 602	10 005	108 010	108 068	96.42	102.30
82·5 . . .	80-84	2 201	0.40709	1.1618	0.47296	527.04	11 597	6 112	57 985	58 022	163.49	154.48
87.5 . . .	85 and over	896					5 485		27 425	27 443	259.56	259.56

(a) This refers to the crude female birth rate $b_f = 45.075$ per thousand.
(b) The first figure in the column represents the ratio C_0/b, the second $C_{1\text{-}4}/4C_0$ and the third $4C_{5\text{-}9}/5C_{1\text{-}4}$.
(c) The first figure in the column corresponds to $k = 0.5$, the second to $k = 2.5$, the third to $k = 4.5$ and all the others to $k = 5$.

TABLE IV.6. COMPARISON OF DEATH RATES BY AGE GROUPS m_a IN TWO MALTHUSIAN POPULATIONS COMPUTED ON THE BASIS OF AGE DISTRIBUTION IN TABLE II.3 (ASSUMED TO BE CONSTANT) BY TAKING SUCCESSIVELY TWO VALUES, 0.030 AND 0.015, AS THE RATE OF NATURAL VARIATION r

Age group (a)	Death rate m_a (per thousand)		Difference
	$r = 0,015$ (a)	$r = 0,030$ (b)	
0	147.53	136.41	11.12
1-4	33.53	16.19	17.34
5-9	17.56	3.72	13.84
10-14	18.52	3.57	14.95
15-19	19.36	4.40	14.96
20-24	20.76	5.67	15.09
25-29	21.43	6.41	15.02
30-34	21.86	6.83	15.03
35-39	22.39	7.36	15.03
40-44	23.37	8.37	15.00
45-49	25.19	10.28	14.91
50-54	28.38	13.45	14.93
55-59	33.52	18.59	14.93
60-64	41.98	27.15	14.83
65-69	55.90	41.12	14.78
70-74	78.00	63.08	14.92
75-79	111.97	96.42	15.55
80-84	169.68	163.49	6.19
85 and over .	298.97	259.56	39.41

(a) Rate taken from column 8 of table IV.2.
(b) Rate taken from column 12 of table IV.5.

TABLE IV.7. AGE DISTRIBUTION OF THE FEMALE POPULATION OF BRAZIL AT THE 1900, 1940 AND 1950 CENSUSES, AND ADJUSTED DISTRIBUTION BASED ON THE RESULTS OF THE CENSUSES

Age group (years)	Year of census			Adjusted distribution
	1900	1940	1950	
0-4	4 443	1 546	1 590	1 764
5-9		1 376	1 329	1 367
10-14		1 284	1 209	1 203
15-19	1 134	1 110	1 099	1 047
20-24	927	960	1 003	902
25-29	864	829	808	776
30-34	532	622	624	650
35-39	624	560	583	544
40-44	356	459	447	456
45-49	387	343	368	378
50-54	195	294	298	301
55-59	223	187	198	223
60-64	101	171	178	155
65-69	94	97	100	107
70-74	40	76	76	58
75-79	41	38	39	39
80-84	15	27	28	20
85 and over .	24	21	23	10
ALL AGES . .	10 000	10 000	10 000	10 000
0-14	4 443	4 206	4 128	4 334

adjusted to the female population of Brazil, as recorded in the censuses of 1900, 1940 and 1950. The age distributions of these three censuses are close to one another and the adjustment, made graphically by hand, simply smoothed out some irregularities in the age pyramid. Table IV.7 gives the three age distributions provided by the censuses as will as the smooth and adjusted age distribution. It should be noted in passing, however, that there is still a good deal of uncertainty about the first age group. Moreover, it should be noted that this first age group covers the first five years of life, whereas in the preceding example those five years were divided into the age groups under 1 year and 1-4.

Let us now consider the sub-set F(r) attached to the adjusted age distribution, and let us determine which of the populations of that sub-set corresponds to $r = 0.025$. We can use either of the two methods already mentioned:

(1) We can determine the sequence of the quantities: b_0L_{0-4}, b_0L_{5-9}, b_0L_{10-14} etc., by multiplying the sequence of the quantities, C_{0-4}, C_{5-9}, C_{10-14}, etc. by $e^{2.5r}$, $e^{7.5r}$, $e^{12.5r}$ etc., respectively. We can then determine without difficulty the series of quantities: $b_0p(5)$, $b_0p(10)$,$b_0p(15)$ etc.

(2) We can also calculate the survivorship rates directly by multiplying the quantities:

$$\frac{C_{5-9}}{C_{0-4}}, \frac{C_{10-14}}{C_{5-9}}, \frac{C_{15-19}}{C_{10-14}}, \text{etc.,}$$

by e^{5r}. We thus obtain the survivorship rates from 2.5 7.5 years of age, from 7.5 to 12.5, from 12.5 to 17.5. and so on.

In both methods it is necessary, in order to complete the computation, to make an assumption regarding b_0, and we must therefore begin by determining the series of birth rates compatible with the age structure. Table IV.8 gives the essentials of the computation.

In order to complete the computation, we have arbitrarily assumed that $b_0 = 45$ per thousand. As can be seen from table IV.8, this is tantamount to assuming an infant mortality of between 96 and 106 per thousand, depending on the value of λ (96 per thousand corresponds to $\lambda - 0.6$, while 106 per thousand corresponds to $\lambda = 0.9$). Once a value has been adopted for b_0, the computation presents no difficulty. Table IV.9 gives details of the computation by the second method, which involves first computing the survivorship rates.

We have dealt with this first example at some length because, as will be seen later, the other examples can always be referred back to it.

SECOND EXAMPLE: MALTHUSIAN POPULATION WITH GIVEN AGE DISTRIBUTION $C_0(a)$ AND KNOWN SURVIVAL RATIO AT A GIVEN AGE a_0

If r_0 represents the rate of natural variation of the population which is sought, we have:

$$p(a_0) = \frac{C_0\,(a_0)}{C_0(0)}\, e^{r_0 a}$$

This formula enables r_0 to be calculated, and this brings us back to the preceding case. In discontinuous notation, we apply the preceding approximate formulae, the best method being to begin by calculating $C_0(a_0)$.

If, for example, $a_0 = 21$ years, we write:

$$C_{15-19} = 5C(17.5)$$

$$C_{20-24} = 5C(22.5)$$

and we determine C(21) by interpolation. We then have the formula:

$$p(21) = \frac{C(21)}{C(0)}\, e^{21 r_0}$$

TABLE IV.8. COMPUTATION OF THE CRUDE BIRTH RATES COMPATIBLE WITH THE AGE STRUCTURE OF THE FEMALE POPULATION OF BRAZIL, AS ADJUSTED TO FIT THE RESULTS OF THE CENSUSES OF 1900, 1940 AND 1950

Median age, a	*Age group*	C_a	e^{-ra}	$C_a e^{ra} = b_0 L_a$	$\frac{b_0 L_a + b_0 L_{a+5}}{10} = b_0\, p(a)$	
			(A) $r = 0.025$			
2.5	0-4	1 764	0.93941	1 878		
7.5	5-9	1 367	0.82901	1 649	330.3 (a)	$b_0 p(5)$
12.5	10-14	1 203	0.73162	1 644	329.3	$b_0 p(10)$
17.5	15-19	1 047	0.64565	1 622	326.6	$b_0 p(15)$
			(B) $r = 0$ (stationary population)			
2.5	0-4	1 764		1 764		
7.5	5-9	1 367		1 367	289.8 (b)	$b_0 p(5)$
12.5	10-14	1 203		1 203	257.0	$b_0 p(10)$
17.5	15-19	1 047		1 047	225.0	$b_0 p(15)$

(C) Compatible crude birth rates (per thousand), application of formula IV.4

	Value of λ:			
	0.6		*0.9*	
Infant mortality (q_0 per thousand)	*$r = 0.025$*	*$r = 0$*	*$r = 0.025$*	*$r = 0$*
10	41.2	40.2	41.2	40.2
50	42.7	41.6	42.9	41.9
100	44.7	43.6	45.2	44.1
200	49.3	48.1	50.6	49.4
300	55.1	53.7	57.5	56.1
400	62.3	60.8	66.5	64.9
500	71.8	70.0	78.9	77.0

(a) We assume that $b_0 p(7.5) = 1\,649/5 = 329.8$ and we selected $b_0 p(5)$ symmetric of $b_0 p(10)$ with respect to $b_0 p(7.5)$.

(b) We assume that $b_0 p(7.5) = 1\,367/5 = 273.4$ and we selected $b_0 p(5)$ symmetric of $b_0 p(10)$ with respect to $b_0 p(7.5)$.

which makes it possible to calculate r_0, provided that C(0)—in other words, b_0—is known. If it is not known, it can be chosen from the series of birth rates compatible with the age distribution. There is thus an infinity of possible values of r_0 and consequently an infinity of populations satisfying the given conditions. Thus, the problem, which was a determinate one in continuous notation, becomes indeterminate in discontinuous notation.

THIRD EXAMPLE: MALTHUSIAN POPULATION WITH GIVEN AGE DISTRIBUTION $C_0(a)$ AND KNOWN CRUDE DEATH RATE d_0

In continuous notation, the problem is determinate, since knowledge of d_0 is tantamount to knowledge of $r_0 = b_0 - d_0 = C(0) - d_0$, thus bringing us back to the first example. In discontinuous notation, however, the problem is not determinate. All values of r_0, such as $r_0 = b_0 - d_0$, are valid, provided that b_0 is within the range of values compatible with the age distribution.

FOURTH EXAMPLE: MALTHUSIAN POPULATION WITH GIVEN AGE DISTRIBUTION $C_0(a)$ AND KNOWN AGE DISTRIBUTION OF DEATHS AT A GIVEN AGE a_0

In continuous notation, we calculate r by the formula:

$$r = \frac{C(0)d(a_0) + C'(a_0)}{-C(a_0) + d(a_0)} \qquad \text{(II.12)}$$

In discontinuous notation—if, for example, we know the proportion $d_{20\text{-}24}$ of deaths in the 20-24 age group—we shall assume that we can write:

$$5d(22.5) = d_{20\text{-}24}$$

and

$$5C(22.5) = C_{20\text{-}24}$$

In order to apply the preceding formula, we must calculate C′(22.5). We assume that we can write:

$$5C'(22{,}5) = 5\,\frac{C(17.5) - C(27.5)}{10} =$$

$$= \frac{1}{2}\,\frac{C_{15\text{-}19}}{5} - \frac{C_{25\text{-}29}}{5} = \frac{1}{10}(C_{15\text{-}19} - C_{25\text{-}29})$$

TABLE IV.9. COMPUTATION OF THE FEMALE MORTALITY RATES IN A MALTHUSIAN FEMALE POPULATION WITH A DISTRIBUTION BY AGE GROUPS ADJUSTED TO FIT THE FEMALE POPULATION OF BRAZIL, AS RECORDED IN THE CENSUSES OF 1900, 1940 AND 1950, AND WITH A RATE OF NATURAL VARIATION OF $r = 0.025$

Median age, a	*Age group (years)* a	*Distribution by age groups* C_a	$\frac{C_{a+5}}{C_a}$	*Multiplier* (a) $k = e^{5r}$	$k\frac{C_{a+5}}{C_a}$	*Probability of death from one age to the next (per 1 000)*	*Survivors to age* a	*Deaths between age* a *and the next age*	*Survivors in each age group*	*Survivors to initial age of age group*	*Deaths between one age and the next*	*Death rate* m_a *(per 1 000)*
0	Births (5b)	2 250	0.7840	1.0645 (a)	0.8346	165.4	10 000	1 654				
2.5 . . .	0-4	1 764	0.7749	1.1332	0.8781	121.9	8 346	1 017	41 730	10 000	2 660	63.7
7.5 . . .	5-9	1 367	0.8800	1.1332	0.9972	2.8	7 329	21	36 645	7 340	22	0.6
12.5 . . .	10-14	1 203	0.8703	1.1332	0.9862	13.8	7 308	101	36 540	7 318	60	1.7
17.5 . . .	15-19	1 047	0.8615	1.1332	0.9762	23.8	7 207	171	36 035	7 258	137	3.8
22.5 . . .	20-24	902	0.8603	1.1332	0.9748	25.2	7 036	177	35 180	7 121	173	4.9
27.5 . . .	25-29	776	0.8376	1.1332	0.9791	50.9	6 859	349	34 295	6 948	264	7.7
32.5 . . .	30-34	650	0.8369	1.1332	0.9483	51.7	6 510	336	32 550	6 684	342	10.5
37.5 . . .	35-39	544	0.8382	1.1332	0.9498	50.2	6 174	310	30 870	6 342	323	10.5
42.5 . . .	40-44	456	0.8289	1.1332	0.9393	60.7	5 864	356	29 320	6 019	333	11.3
47.5 . . .	45-49	378	0.7963	1.1332	0.9023	97.7	5 508	537	27 540	5 686	446	16.2
52.5 . . .	50-54	301	0.7409	1.1332	0.8396	160.4	4 971	798	24 855	5 240	668	26.9
57.5 . . .	55-59	223	0.6951	1.1332	0.7877	212.3	4 173	886	20 865	4 572	842	40.4
62.5 . . .	60-64	155	0.6903	1.1332	0.7822	217.8	3 287	718	16 435	3 730	802	48.8
67.5 . . .	65-69	107	0.5421	1.1332	0.6143	305.7	2 569	981	12 845	2 928	850	66.1
72.5 . . .	70-74	58	0.6724	1.1332	0.7619	238.1	1 588	385	7 940	2 078	682	86.0
77.5 . . .	75-79	39	0.5128	1.1332	0.5811	418.9	1 203	504	6 015	1 396	445	73.9
82.5 . . .	80-84	20	0.5000	1.1332	0.5666	433.4	699	303	3 495	951	404	115.5
87.5 . . .	85 +	10					396		1 980	547	547	276.5

(a) Except for the first line, where $k = e^{2.5r}$

We shall finally obtain:

$$r = \frac{b_0 d_{20\text{-}24} + \frac{1}{10}(C_{15\text{-}19} - C_{25\text{-}29})}{-C_{20\text{-}24} + d_{20\text{-}24}} \qquad \text{(IV.5)}$$

As in the previous examples, to each value of r corresponds some value of b_0 which is compatible with the age distribution. There will thus be an infinite number of populations satisfying the condition.

Finally, in discontinuous notation, formula IV.5, which gives the value of r by approximation, is not valid at the beginning and the end of life.

Fifth example: Malthusian population with given age distribution $C_0(a)$ and known probability of dying at a given age $q(a_0)$

In continuous notation we write:

$$q(a_0) + \frac{C'(a_0)}{C(a_0)} = -r$$

which brings us back to the first example.

In discontinuous notation, the formulae which we have written for the fourth example make it easy for us to calculate such quantities as:

$$\frac{C'(22.5)}{C(22.5)} = \frac{C_{15\text{-}19} - C_{25\text{-}29}}{10C_{20\text{-}24}}$$

Furthermore, if $m_{20\text{-}24}$ is the death rate between the ages of 20 and 24, we have, approximately, $q(22.5) = m_{20\text{-}24}$, whence we have the formula;

$$m_{20\text{-}24} + \frac{C_{15\text{-}19} - C_{25\text{-}29}}{10C_{20\text{-}24}} + r = 0$$

Sixth example: Malthusian population with given age distribution $C_0(a)$ and known mean death rate for five-year groups between the ages of 5 and 34

In this example, the quantity which we assume to be known in addition to the age distribution is written:

$$A = \frac{(m_{5\text{-}9} + m_{10\text{-}14} + m_{15\text{-}19} + m_{20\text{-}24} + m_{25\text{-}29} + m_{30\text{-}34})}{6}$$

In accordance with the formula in the fifth example, we can write:

$$A = -r - \frac{1}{60}\left[\frac{C_{0\text{-}4} - C_{10\text{-}14}}{C_{5\text{-}9}} + \frac{C_{5\text{-}9} - C_{15\text{-}19}}{C_{10\text{-}14}} + \dots + \frac{C_{25\text{-}29} - C_{35\text{-}39}}{C_{30\text{-}34}}\right],$$

This formula enables us to calculate r, thus bringing us back to the first example.

Introduction of statistical variables

Although it would be easy to think of many other examples, we shall confine ourselves to the six described above. All these examples lead us to an r equation the solution of which gives us an *exact and determinate* value of r or, at most, a small number of exact and determinate values. If we are satisfied with approximate values, then we must resolve equations whose solutions are "statistical variables" diverging to a greater or lesser degree from the mean values, and new problems then arise. We have already encountered problems of this kind when dealing with the sub-sets $H(r)$, and we have already studied the compatibility between:

(*a*) An age distribution of the population and a given death rate;

(*b*) An age distribution of deaths and a given death rate.

In the study of the sub-set $F_0(r)$ a third problem is encountered: that of the *compatibility of a given age distribution of deaths and a given age distribution of the population.*

In a sub-set $F_0(r)$ we cannot arbitrarily assume knowledge of the entire age distribution of deaths. As was seen in the fourth example above, knowledge of this age distribution of deaths at a single age $d(a_0)$ was sufficient to determine a population from $F_0(r)$ and was therefore sufficient to define $d(a)$ at all ages.

Let us now consider an actual population in which $C_0(a)$ and $d_0(a)$ are the observed age distributions of the population and the distribution of the observed deaths. Let us take the sub-set $F_0(r)$ corresponding to $C_0(a)$ and pose the following question: is the age distribution $d_0(a)$ consistent with the sub-set $F_0(r)$? If we want $d_0(a)$ to coincide *exactly* with the age distribution of deaths of a population from $F_0(r)$, then obviously the answer will in most cases be "no", for it is unlikely that we should find in $F_0(r)$ an age distribution of deaths which coincides exactly with the observed age distribution $d_0(a)$. If we are satisfied with an approximate coincidence, however, the question takes on a different aspect.

If there is exact coincidence, we can write in accordance with the formulae in table II.2 (see chapter II):

$$-\frac{C'(a)}{C(a)} - r = \frac{d(a)}{C(a)}(b_0 - r)$$

which is written:

$$C'(a) + rC(a) + dd(a) = 0$$

If we put:

$$\frac{C'(a)}{d(a)} = y \quad \text{and} \quad \frac{C(a)}{d(a)} = x$$

we therefore have:

$$y + rx + d = 0 \qquad \text{(IV.6)}$$

For each value of a we have a pair of values (x, y), and if we plot a graph with x on the horizontal axis and y on the vertical axis, the points obtained will be on a straight line defined by the equation IV.6.

If there is no exact coincidence, we shall obtain a cluster of points to which a straight line can be fitted and whose equation enables us to compute r_0. The population from $F_0(r)$ corresponding to r_0, which is the population sought, is one whose age distribution of deaths *coincides approximately* with the age distribution of deaths $d_0(a)$ actually observed.

In discontinuous notation, the formulae giving x and y are easily written. In accordance with the previous example, we have:

$$5C'(22.5) = \frac{1}{10}(C_{15\text{-}19} - C_{25\text{-}29})$$

$$5\,C(22.5) = C_{20\text{-}24}$$

and $\quad 5d(22.5) = d_{20\text{-}24}$

We therefore have:

$$y = \frac{C_{15\text{-}19} - C_{25\text{-}29}}{10 d_{20\text{-}24}}\, y$$

$$x = \frac{C_{20\text{-}24}}{d_{20\text{-}24}}\, x$$

Tables IV.11 and IV.12 give an example of the application of this formula, using for C_a the age distribution adjusted to fit the results of the population censuses carried out in Mexico in 1930, 1940 and 1950 and for d_a the age distribution of deaths actually observed in Mexico in 1950.

Table IV.10 shows that the age distribution of the female population of Mexico varied only slightly in the three censuses mentioned. Graph IV.1 shows how the adjustment was made. It was drawn by hand so as to eliminate the fluctuations in age distribution observed in the three censuses. As a result, the fluctuations caused by variations in the birth rate during the civil war of 1911-1921 and the period which followed it were eliminated.

In the case of the female deaths observed in 1950, we have taken the crude figures without making any adjustment. This does not mean that we considered the crude figures to be more accurate than the population figures given by the censuses; on the contrary, the age distribution

TABLE IV.10. AGE DISTRIBUTION OF THE FEMALE POPULATION OF MEXICO, AS RECORDED IN THE 1930, 1940 AND 1950 CENSUSES, AND AGE DISTRIBUTION ADJUSTED ON THE BASIS OF THE RESULTS OF THE THREE CENSUSES

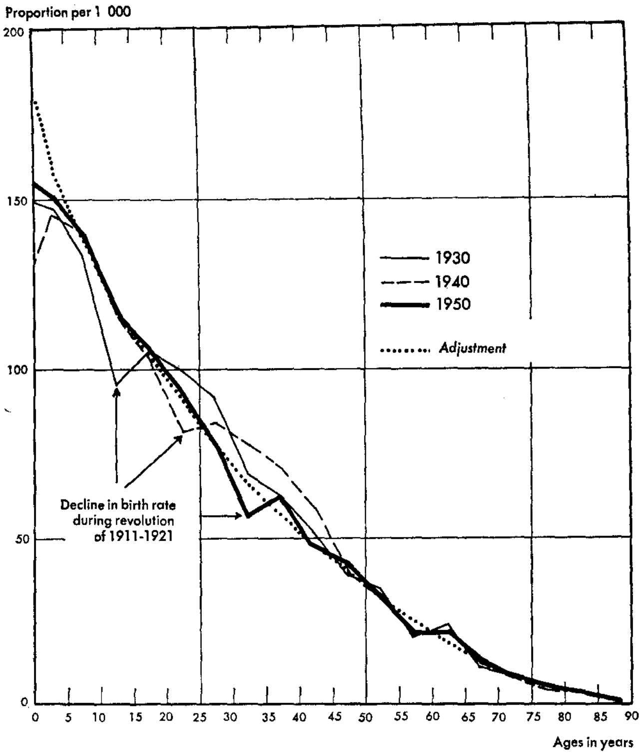
Age group (years)	*Year of census* 1930	1940	1950	*Adjusted distribution*
Under 1 year	29 772	26 278	30 830	35 796
1-4.	117 442	115 992	119 929	125 921
5-9	133 338	139 361	138 465	135 521
10-14	95 335 (a)	116 107	115 558	117 193
15-19	105 799 (b)	103 136	105 875	103 624
20-24	99 863	81 141 (a)	94 320	90 714
25-29	91 695	84 317 (b)	79 433	77 745
30-34	68 903	68 743	56 082 (a)	65 707
35-39	62 710	70 407	61 097 (b)	55 900
40-44	50 661	48 971	47 634	46 621
45-49	38 067	39 699	41 235	40 358
50-54	34 326	31 818	32 359	31 672
55-59	19 261	22 054	20 411	24 027
60-64	23 547	21 571	22 106	18 633
65-69	10 175	11 583	12 967	12 691
70-74	8 866	8 446	9 735	9 528
75-79	4 014	4 500	5 049	4 942
80-84	3 831	3 363	3 941	2 603
85-89	1 139			
90-94	734	2 510	2 974	804
95-99	336			
100 and over.	187			
ALL AGES . .	1 000 000	1 000 000	1 000 000	1 000 000

(a) Decline in birth rate during revolution of 1911-1921.
(b) Resurgence of birth rate after the civil war.

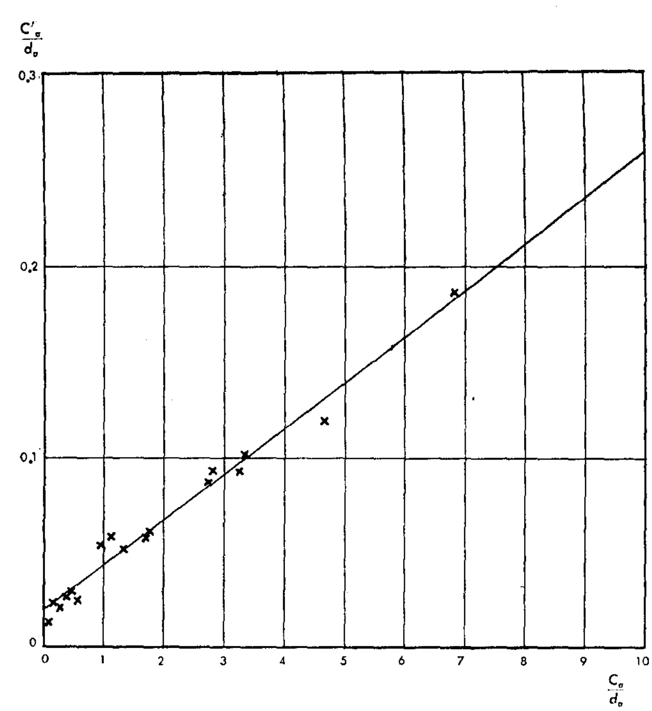
TABLE IV.11. COMPUTATION OF THE DERIVATIVE $C'(a)$ CORRESPONDING TO THE AGE DISTRIBUTION OF THE FEMALE POPULATION OF MEXICO, AS ADJUSTED TO FIT THE RESULTS OF THE CENSUSES OF 1930, 1940 AND 1950 (SEE TABLE IV.10)

Median age a	*Age group (years)* a	*Distribution by age groups* (a) C_a	*Distribution by years of age for median years* $C(a) = C_a/5$	*Successive differences* $C(a) - C(a+5)$	*Average of two successive differences* $C'(a)$
0.5	0	35 796	35 796 (b)	1 726.4 (d)	
3.0	1-4	125 921	31 480 (c)	972.6 (e)	1 349.5 (f)
7.5	5-9	135 521	27 104	733.0	852.8 (f)
12.5	10-14	117 193	23 439	542.8	637.9
17.5	15-19	103 624	20 725	516.4	529.6
22.5	20-24	90 714	18 143	518.8	517.6
27.5	25-29	77 745	15 549	481.6	500.2
32.5	30-34	65 707	13 141	392.2	436.9
37.5	35-39	55 900	11 180	371.2	381.7
42.5	40-44	46 621	9 324	250.4	310.8
47.5	45-49	40 358	8 072	347.6	299.0
52.5	50-54	31 672	6 334	305.8	326.7
57.5	55-59	24 027	4 805	215.6	260.7
62.5	60-64	18 633	3 727	237.8	226.7
67.5	65-69	12 691	2 538	126.4	182.1
72.5	70-74	9 528	1 906	183.6	155.0
77.5	75-79	4 942	988	93.4	138.5
82.5	80-84	2 603	521	72.0	82.7 (f)
87.5	85 and over	804	161		
	ALL AGES	1 000 000			

(a) Adjusted age distribution of table IV.10.
(b) $C(0.5) = C_0$.
(c) $C(3.0) = (1/4)C_{1-4}$.
(d) $C(0.5) - C(3.0)/2.5$.
(e) $C(3.0) - C(7.5)/4.5$.
(f) Approximate figures.



Graph IV.1. Distribution by five-year age groups of the female population of Mexico, as recorded in the 1930, 1940 and 1950 censuses— and an adjusted distribution on the basis of the results of those three censuses



Graph IV.2. Compatibility of the age distribution of the female population of Mexico, as adjusted to fit the age distributions recorded in the 1930, 1940 and 1950 censuses—with the age distribution of female deaths recorded in Mexico in 1950

TABLE IV.12. COMPATIBILITY OF THE AGE DISTRIBUTION OF A POPULATION AND ITS AGE DISTRIBUTION OF DEATHS

Computation of the quantities x and y of equation IV.6 in the text, using an age distribution adjusted to that of the female population of Mexico, as recorded in the censuses of 1930, 1940 and 1950 (see table IV.10), and the age distribution of female deaths recorded in Mexico in 1950

Median age a	*Age group (years)* a	C(a) *Figures from column 4 of table IV.11*	C′(a) *Figures from the last column of table IV.11*	*Distribution by age group of female deaths* (a) d_a	*Distribution by years of age of female deaths, for the median age* $d(a) = d_a/5$	$\frac{C(a)}{d(a)} = x$	$\frac{C'(a)}{d(a)} = y$
0.5	0	35 796		256 964	256 964 (b)		
3.0	1-4	31 480	1 349.5 (c)	221 823	55 431 (d)	0.568	0.0243 (c)
7.5	5-9	27 104	852.8 (c)	40 466	8 093	3.349	0.1053 (c)
12.5	10-14	23 439	637.9	17 141	3 428	6.837	0.1861
17.5	15-19	20 725	529.6	22 245	4 449	4.658	0.1195
22.5	20-24	18 143	517.6	27 877	5 575	3.254	0.0928
27.5	25-29	15 549	500.2	28 168	5,634	2.760	0.0888
32.5	30-34	13 141	436.9	23 396	4 679	2.809	0.0934
37.5	35-39	11 180	381.7	31 107	6 221	1.797	0.0614
42.5	40-44	9 324	310.8	26 139	5 228	1.784	0.0595
47.5	45-49	8 072	299.0	28 861	5 772	1.330	0.0518
52.5	50-54	6 334	326.7	27 922	5 584	1.134	0.0585
57.5	55-59	4 805	260.7	24 320	4 864	0.988	0.0537
62.5	60-64	3 727	226.7	38 411	7 682	0.485	0.0295
67.5	65-69	2 538	182.1	35 025	7 005	0.362	0.0260
72.5	70-74	1 906	155.0	38 391	7 678	0.248	0.0202
77.5	75-79	988	138.5	29 329	5 866	0.168	0.0236
82.5	80-84	521	82.7 (c)	31 524	6 305	0.083	0.0131 (c)
87.5	85 and over	161		50 890	10 178		
	ALL AGES			1 000 000			

SOURCE: United Nations *Demographic Yearbooks.*

(a) Female deaths recorded in Mexico in 1950.

(b) $d(0.5) = d_0$.

(c) Approximate figures.

(d) $d(3.0) = (1/4)d_{1-4}$.

of deaths is no doubt less accurate than the age distribution of the population—but the errors involved are such as to be very difficult to correct and, since any adjustment would be arbitrary, it was decided to keep the crude figures as they were.

Graph IV.2 was drawn by using the values x and y of table IV.12. As may be seen, to the points obtained can easily be fitted a straight line with the equation:

$$y + rx + d = 0$$

The ordinate-intercept of this straight line is equal to d. Graph IV.2 gives: $d = 0.020$.

The slope of the straight line is equal to r and we have: $r = 0.024$, whence we obtain the crude birth rate $b = d + r = 0.044$.

Thus, the sub-set $F_0(r)$ of Malthusian populations associated with the adjusted age distribution of the female population of Mexico, as recorded in the 1930, 1940 and 1950 censuses and having the vital rates:

$$\begin{cases} r = 0.024 \\ d = 0.020 \\ b = 0.044 \end{cases}$$

has an age distribution of female deaths which coincides *approximately* with the age distribution of female deaths observed in Mexico in 1950.

We shall now leave the sub-sets $F(r)$ of Malthusian populations and turn to an examination of the "processes of demographic evolution where the age distribution is invariable".