Arbeidsnotater

STATISTISK SENTRALBYRÅ

IO 68/6 Oslo, 3. mai 1968

Fertility rates and reproduction rates in a probabilistic setting.

by Jan M. Hoem

Contents

Summary: In the setting of a very simple probabilistic fertility model, net and gross reproduction rates are defined and the connection between their "continuous" and "discrete" age versions is brought out. It is shown that the interpretations of the common definitions of the reproduction rates are inconsistent, but that the numerical effect on the rates is probably negligible in most cases. Maximum likelihood and some other estimators of parameters of the model are set up and are compared with methods in common use.

1. Introduction

§ 1.1. In common texts on demography, the net reproduction rate is generally defined as

$$R_N = \int_{\omega_1}^{\omega_2} \frac{\ell_x}{\ell_0} \cdot f_x \, dx, \tag{1.1}$$

where $\{\ell_x\}$ is a decrement series, $\{f_x\}$ is a series of age-specific fertility rates, and the ages between ω_1 and ω_2 constistute the fertile period. We shall call this a continuous age version of the definition. In addition, one may find the discrete age version

$$R_N = \sum_{x=\omega_1}^{\omega_2 - 1} \frac{\ell_x}{\ell_0} f_x, \tag{1.2}$$

which is commonly used for purposes of numerical calculation.

Similarly there is a continuous and a discrete age version of the definition of the gross reproduction rate:

$$R_B = \int_{\omega_1}^{\omega_2} f_x \, dx \quad \text{and} \quad R_B = \sum_{x=\omega_1}^{\omega_2-1} f_x. \tag{1.3}$$

The fertility rate f_x may be specified for a female or for a male parent, and fertility may be defined with respect to live male children, to live female children, or to live children of both sexes. To each of these specifications of the fertility rate (and a consistent specification of the decrement series) there corresponds a net and a gross reproduction rate.

It need not concern us here exactly which specification has been made, nor does it matter whether decrement series and fertility rates are calculated on a calendar or a generation basis. What we have to say, will be relevant to any of these alternatives.

When we speak of parent and child, it will be understood that there is a specified sex for the parent and a specified kind of live child. By births we shall mean live births.

§ 1.2. The accepted interpretation of the reproduction rates may be formulated as follows:

R_N is the average number of children that will be born to a parent, evaluated at the birth of the parent. (1.4)

R_B is the same average number of children, provided the parent survives the fertile period, i. e. provided there is no mortality before age ω_2. (1.5)

The averages are evaluated on the assumption that the mortality (in the case of R_N) and the fertility (in the case of both R_N and R_B) of the parent develops as specified by the series $\{\ell_x\}$ and $\{f_x\}$.

§ 1.3. Several questions should have arisen up to this point:

(i) How are the two versions of the definition of each reproduction rate reconciled ?

(ii) Are (say) (1.2) and the second member of (1.3) consistent ?

(iii) Is the formula for a reproduction rate consistent with its verbal interpretation, as given in § 1.2 ?

One purpose of the present paper is to answer these questions. We shall introduce a very simple probabilistic model which will enable us to answer question (i) directly. Within this model, which gives a reasonable interpretation of the calculations made in connection with the reproduction rates, we shall give the following answer to questions (ii) and (iii):

Suppose that we make (1.1) and (1.2) consistent with (1.4). Then the second member of (1.3) is inconsistent with (1.5). And vice versa.

We shall show, however, that in a certain sense the numerical effect of the inconsistency on the reproduction rates is probably negligible in most cases.

Finally, we shall set up maximum likelihood and some other estimators for the parameters of the model, and shall see how these may be applied in practice.

§ 1.4. Essentially the same model as the one presented here has been briefly considered previously by Joshi (1954), and by Consaël and Lamens (1962).

2. The model

§ 2.1. We shall let $\mu(x)$ be the force of mortality and $\phi(x)$ be the force of fertility for an x year old parent. This will be taken to mean the following:

Let us observe a parent alive at age x during the age interval $\langle x, x + \Delta x \rangle$ with $\Delta x > 0$. Then

(i) the probability that the parent will die in the age interval without giving birth to any children, equals $\mu(x)\Delta x + o(\Delta x)$, where $o(\Delta x)/\Delta x \to 0$ as $\Delta x \to 0$,

(ii) the probability that the parent will have exactly one birth in the age interval and survive to age $x + \Delta x$, equals $\phi(x)\Delta x + o(\Delta x)$,

(iii) the probability that the parent will survive to age $x + \Delta x$ without giving birth to any children in the age interval, equals $1 - [\mu(x) + \phi(x)]\, \Delta x + o(\Delta x)$, and

(iv) the probability that the parent gives more than one birth, or gives at least one birth and then dies within the age interval, is $o(\Delta x)$.

We shall assume that $\mu(.)$ and $\phi(.)$ are continuous functions for $x \geqq 0$, with $\mu(x) > 0$ for $0 \leqq x < \omega$, $\phi(x) > 0$ for $\omega_1 < x < \omega_2$, $\phi(x) = 0$ otherwise. Thus the fertile period is the period where $\phi(x) > 0$. Multiple births will be taken care of at a later stage. (See § 2.7.)

§ 2.2. It will be seen that in the present model, fertility does not depend on the marital status of the parent, parity of the birth, time elapsed since last birth, etc. We do not claim that this model gives a realistic description of all major aspects of true-life fertility in any sense. We do believe that it has some of the essential features of fertility, however, and what is even more important: it permits us to give a precise interpretation of many of the measures commonly used in fertility and reproductivity analysis. It is our belief that such measures have little meaning except within some model. It can then do no harm to formulate it explicitly, and the formulation permits us to make such important distinctions as the one between a parameter, estimators for it, and estimates calculated. (For the terminology, see e.g. Kendall and Buckland (1957)).

One further motive for choosing the model of § 2.1 is its simplicity, which will permit us to make our points without expending too large an effort on the mathematics of the problem.

§ 2.3. We shall need the following quantities:

$$_tp_x = \frac{\ell_{x+t}}{\ell_x} = \exp\{-\int_0^t \mu(x+\tau)\,d\tau\}$$

is the probability that a parent of age x will survive to age x + t.

$$_tq_x = 1 - {}_tp_x.$$

$P_k(x,t)$ is the probability that a parent of age x will give birth to k children in the age interval $[x, x+t>$ and will survive to age x + t.

$Q_k(x,t)$ is the probability that a parent of age x will give birth to k children in the age interval $[x, x+t>$ and will die in the interval.

$$R_k(x,t) = P_k(x,t) + Q_k(x,t) \tag{2.1}$$

is the probability that a parent of age x will give birth to k children in the age interval $[x, x+t>$.

$$f_1(x,t) = \sum_{k \geq 1} k\, P_k(x,t) \,/\, {}_tp_x \tag{2.2}$$

is the expected number of children to which a parent of age x will give birth in the age interval $[x, x+t\rangle$, given that the parent survives the interval.

$$f_2(x,t) = \sum_{k \geq 1} k\, Q_k(x,t) \,/\, {}_tq_x$$

is the corresponding expected number of children, given that the parent does not survive the interval.

$$f(x,t) = \sum_{k \geq 1} k\, R_k(x,t) = {}_tp_x\, f_1(x,t) + {}_tq_x\, f_2(x,t) \tag{2.3}$$

is the expected number of children to which a parent will give birth in the age interval $[x, x+t\rangle$.

§ 2.4. Let us first consider $P_k(x,t)$ and $f_1(x,t)$ for $x < \omega_2$. To assume that it is given that the parent will survive to age x+t, amounts to making evaluations as if $\mu(x+\tau) = 0$ for $0 \leq \tau \leq t$. (A formal proof will be given in chapter 8.) In this case the births in the interval are generated by a Poisson process with variable intensity $\phi(x+\tau)$ for $0 \leq \tau \leq t$. As shown e.g. by Khinchine (1960, pp. 18-21) we then have

$$P_k(x,t) \,/\, {}_tp_x = \frac{1}{k!} \left\{ \int_0^t \phi(x+\tau)\, d\tau \right\}^k \exp\left\{ -\int_0^t \phi(x+\tau)\, d\tau \right\}$$

for $k = 0, 1, 2, \ldots$. Therefore,

$$f_1(x,t) = \int_0^t \phi(x+\tau)\, d\tau \tag{2.4}$$

and
$$P_k(x,t) = \frac{1}{k!} \left[f_1(x,t) \right]^k \exp\left\{ -\int_0^t \left[\mu(x+\tau) + \phi(x+\tau) \right] d\tau \right\} \tag{2.5}$$

for $k = 0, 1, 2, \ldots$.

Summation over k in (2.5) gives

$$ {}_tp_x = \sum_{k=0}^{\infty} P_k(x,t) \tag{2.6}$$

as we ought to get.

§ 2.5. By decomposing according to the moment of the last birth in the age interval $[x, x + t>$ we obtain for $k \geq 1$

$$P_k(x,t) = \int_0^t P_{k-1}(x,\tau)\, \phi(x + \tau)\, P_0(x + \tau, t - \tau)\, d\tau, \tag{2.7}$$

$$Q_k(x,t) = \int_0^t P_{k-1}(x,\tau)\, \phi(x + \tau)\, Q_0(x + \tau, t - \tau)\, d\tau, \tag{2.8}$$

and consequently

$$R_k(x,t) = \int_0^t P_{k-1}(x,\tau)\, \phi(x + \tau)\, R_0(x + \tau, t - \tau)\, d\tau. \tag{2.9}$$

If we multiply by k in (2.9) and add for all $k \geq 1$, we obtain by the Lebesgue monotone convergence theorem

$$f(x,t) = \int_0^t \sum_{k \geq 0} (k+1)\, P_k(x,\tau)\, \phi(x+t)\, P_0(x+\tau, t-\tau)\, d\tau.$$

By (2.2) and (2.6) this gives

$$f(x,t) = \int_0^t {}_\tau p_x \left[f_1(x,\tau) + 1\right] \phi(x + t)\, P_0(x + \tau, t - \tau)\, d\tau,$$

which can be interpreted directly. Starting from (2.7) or (2.8) we may obtain quite similar formulae for $f_1(x,t)$ and $f_2(x,t)$.

§ 2.6. Since obviously for $\Delta t > 0$:

$$Q_k(x, t + \Delta t) = Q_k(x,t) + P_k(x,t) \,.\, \mu(x + t)\, \Delta t + o(\Delta t),$$

we get

$$\frac{\partial}{\partial t} Q_k(x,t) = P_k(x,t)\, \mu(x + t),$$

and

$$Q_k(x,t) = \int_0^t P_k(x,\tau)\, \mu(x + \tau)\, d\tau. \tag{2.10}$$

By (2.5), $\frac{\partial}{\partial t} P_0(x,t) = -[\mu(x + t) + \phi(x + t)] \ P_0(x,t)$,

so $$P_0(x,t)\mu(x + t) = -\left[\frac{\partial}{\partial t} P_0(x,t) + \phi(x + t) P_0(x,t)\right]. \qquad (2.11)$$

Letting k = 0 in (2.10), introducing (2.11) and making some trivial manipulations, we get

$$R_0(x,t) = 1 - \int_0^t P_0(x,\tau)\phi(x + \tau)d\tau,$$

which can again be interpreted directly.

Multiplying in (2.10) by k and adding, we get

$$f_2(x,t) \cdot {}_tq_x = \int_0^t f_1(x,\tau) \ {}_\tau p_x \ \mu(x + \tau)d\tau$$
$$= \int_0^t f_1(x,\tau)\left[-\frac{\partial}{\partial \tau} {}_\tau p_x\right] d\tau = \Big|_0^t f_1(x,\tau)\left[-{}_\tau p_x\right]$$
$$+ \int_0^t {}_\tau p_x \cdot \frac{\partial}{\partial \tau} f_1(x,\tau)d\tau = - f_1(x,t) \cdot {}_tp_x + \int_0^t {}_\tau p_x\phi(x + \tau)d\tau.$$

Combining this with (2.3), we get

$$f(x,t) = \int_0^t {}_\tau p_x \cdot \phi(x + \tau)d\tau, \qquad (2.12)$$

as we might have guessed. By the mean value theorem,

$$f(x,t) = {}_{\xi_x}p_x \cdot f_1(x,t) \quad \text{for a } \xi_x \epsilon < 0,t>>. \qquad (2.13)$$

In combination with (2.3) this relation gives

$$f_2(x,t) < f(x,t) < f_1(x,t) \qquad (2.14)$$

as long as $<x, x + t > \cap < \omega_1, \omega_2 > \neq \emptyset$, in correspondence with common sense. (It is reasonable to expect that a parent who dies in the age interval $[x, x + t>$ will have fewer births in that interval than a parent who survives age x + t.)

§ 2.7. <u>(Multiple births.)</u> Let B be the actual number of births in the age interval $[x, x + t >$ to a parent of age x. Then B is a random variable, with $P' \{ B = b\} = R_b(x,t)$. Summation in (2.10) gives by (2.6)

$$\sum_{k=0}^{\infty} Q_k(x,t) = \int_0^t {}_\tau p_x \, \mu(x + \tau)d\tau = {}_tq_x .$$

Thus $\sum_{k=0}^{\infty} P' \{B = k\} = \sum_{k=0}^{\infty} P_k(x,t) + \sum_{k=0}^{\infty} Q_k(x,t) = 1$, so $P \{ B < \infty\} = 1$.

If $B > 0$, let C_i be the number of children born in birth no. i. We shall allow for multiple births, and shall let Π_k be the probability that any birth will have the multiplicity k, for $k = 1,2,\dots,K$, with $\sum_{k=1}^{K} \Pi_k = 1$. (We might have introduced stillbirths by permitting $k = 0$, but that is outside the scope of the present paper.) The expected number of children in any birth then equals

$$\Pi = \sum_{k=1}^{K} k \cdot \Pi_k .$$

It may be noted that in this model the Π_k do not depend on the age of the parent.

Given that $B = b > 0$, we shall assume that $C_1, C_2, \dots, C_b$ are independent, each with the distribution $\{\Pi_k\}$. The total number of children born to the parent in the age interval $[x, x + t >$ will be

$$C = \begin{cases} 0 & \text{if } B = 0 , \\ \sum_{i=1}^{B} C_i & \text{if } B > 0 . \end{cases}$$

Then $P' \{C = 0 \} = R_0(x,t)$, and $P \{C = k|B = b\} = \Pi_k^{b*}$ for $b \geq 1$, $k \geq b$,

where the topscript $b*$ signifies the b-th convolution.

Thus $P' \{C = k\} = \sum_{b=1}^{k} \Pi_k^{b*} R_b(x,t)$. Furthermore $E(C|B = b) = b\Pi$, and the expected number of children born in the age interval $[x, x + t >$ to a parent of age x equals $EC = \Pi . f(x,t)$.

A quite similar relation holds if we take it as given that the parent will survive age x+t. We may therefore easily convert an expected number of births into an expected number of born children by multiplication with the factor Π.

3. The reproduction rates

§ 3.1. We define the net reproduction rate as

$$R_N = \Pi f(0, \omega_2), \tag{3.1}$$

and get $$R_N = \int_{\omega_1}^{\omega_2} {}_xp_0 \cdot \bar{\phi}(x)dx \tag{3.2}$$

by (2.12), with $\bar{\phi}(x) = \Pi\phi(x)$. Similarly we define

$$R_B = \Pi f_1(0,\omega_2), \tag{3.3}$$

and (2.4) gives us

$$R_B = \int_{\omega_1}^{\omega_2} \bar{\phi}(x)dx. \tag{3.4}$$

The definitions (3.1) and (3.3) correspond with the verbal definitions (1.4) and (1.5). The formulae (3.2) and (3.4) correspond with (1.1) and the first member of (1.3). We see, therefore, that t h e f_x o f t h e t w o i n t e g r a l s o f § 1.1 m a y b e i n t e r p r e t e d a s a f o r c e o f f e r t i l i t y, m u l t i p l i e d b y t h e e x p e c t e d m u l t i p l i c i t y o f a b i r t h.

Formula (3.2) may be written in the form

$$R_N = \sum_{x=\omega_1}^{\omega_2-1} {}_xp_0 \int_0^1 {}_\tau p_0 \, \bar{\phi}(x + \tau)d\tau.$$

We introduce

$$f_x = \int_0^1 {}_\tau p_x \cdot \bar{\phi}(x + \tau)d\tau = \Pi f(x,1), \tag{3.5}$$

and see that f_x is the expected number of children born in the age interval $[x, x + 1 >$ to a parent of age x. Then

$$R_N = \sum_{x=\omega_1}^{\omega_2-1} {}_xp_0 \cdot f_x, \tag{3.6}$$

which is (1.2), and which corresponds to (1.4).

If we let

$$f'_x = \int_0^1 \bar{\phi}(x + \tau)d\tau = \Pi f_1(x,1), \tag{3.7}$$

then f'_x will be the expected number of children born in the age interval $[x,x+1>$ to a parent of age x who survives to age x + 1. We see that

$$R_B = \sum_{x=\omega_1}^{\omega_2-1} f'_x , \tag{3.8}$$

in correspondence with (1.5) but in contrast to the second member of (1.3).

We see that one should not use the same measure of fertility in the second member of (1.3) as in (1.2). In fact, (2.14) gives $R_B > \sum_{x=\omega_1}^{\omega_2-1} f_x$, and on the other hand $R_N < \sum_{x=\omega_1}^{\omega_2-1} {}_xp_o f'_x$.

§ 3.2. In mortality investigations with several causes of mortality, a distinction is made between partial and influenced probabilities of death[1]. Thus, if there are s causes of mortality and the force of mortality of cause ν is $\mu_x^{(\nu)}$, then

$$q_x^{(\nu)} = \int_0^1 \mu_{x+\tau}^{(\nu)} \exp\{-\int_0^\tau \sum_{\alpha=1}^{s} \mu_{x+\theta}^{(\alpha)} d\theta\} d\tau$$

is the (influenced) probability that a person of age x will die before age x + 1 of cause ν, and

$$q_{x,\nu} = \int_0^1 \mu_{x+\tau}^{(\nu)} \exp\{-\mu_{x+\tau}^{(\nu)}\} d\tau$$

is the corresponding "partial probability". The value of $q_x^{(\nu)}$ is obviously influenced by the forces of mortality of causes $\alpha \neq \nu$, hence the terminology. $q_{x,\nu}$ is the probability that would be effective if all causes other than cause ν were removed.

In analogy with this, one may regard f_x as a measure of <u>influenced fertility</u> at age x, i.e. fertility as influenced by actual mortality. Similarly, f'_x may be taken as a measure of <u>partial fertility</u> at age x, i.e. fertility as it would have been if mortality was inoperative.

---

1) This terminology is due to Sverdrup (1961). In less fortunate but more common terminology, one speaks of independent and dependent probabilities, respectively.

Similarly, R_B may be regarded as a summary measure of partial fertility and R_N as a summary measure of influenced fertility (for all ages taken together).

Behind much of what has been written on the theory of fertility investigation seems to be the notion that one should seek for a "pure" fertility measure cleansed of any influence from mortality. (A particularly clear formulation of this position has been given by Matthiessen (1967, pp. 76-77).) The force $\phi(x)$ of fertility readily points itself out as an excellent measure of this kind. If for pedagogic or other reasons one cannot work with a concept such as the force of fertility, f_x' is an obvious substitute. Since real-life fertility actually is influenced by mortality, I am not at all convinced that one should be content with measuring partial fertility in all cases, however, but would favour employing f_x as well as f_x' to get a fuller picture in cases where mortality is not negligible.

§ 3.3. Formula (2.13) gives

$$f_x = {}_{\xi_x}p_x \cdot f_x' \quad \text{for some} \quad \xi_x \in \langle 0,1 \rangle. \tag{3.9}$$

Using the common approximation $1 - {}_{\xi_x}p_x \approx {}_{\frac{1}{2}}q_x \approx 0{,}5\ q_x$, we get

$$f_x \approx (1 - 0.5 \cdot q_x)\ f_x'. \tag{3.10}$$

In many populations, and certainly in the present Scandinavian ones, $0.5 \cdot q_x$ is so small for all fertile ages (at least for females) that the correction implied by (3.9) is negligible. This is not the case for all ages all over the world, however, and where it is not, (3.10) should be used.

For the net reproduction rate, (3.9) and the mean value theorem for finite sums imply that

$$R_N = \sum_{x=\omega_1}^{\omega_2-1} {}_xp_0 \cdot {}_{\xi_x}p_x \cdot f_x' = {}_{\xi}p_{\eta} \sum_{x=\omega_1}^{\omega_2-1} {}_xp_0 \cdot f_x' \tag{3.11}$$

for some $\xi \in \langle 0,1 \rangle$ and some $\eta \in \langle \omega_1, \omega_2 \rangle$. We may probably use e.g. $0.5 \cdot q_x$

for some x in the range between 25 and 30 as an indicator of the value of ${}_{\xi}q_{\eta}$. Values of this indicator for females in various selected countries are given in table (3.12). The impression is that ${}_{\xi}q_{\eta}$ will be negligible for many populations, but not for all of them.

In a later chapter we shall show that current pratice may be interpreted as consisting in quoting the value of $R_N' = \sum_{x=\omega_1}^{\omega_2-1} {}_xp_0 \cdot f_x'$ for R_N and the value of $\sum_{x=\omega_1}^{\omega_2-1} f_x'$ for R_B. While R_B is therefore correctly estimated, R_N will currently be slightly overestimated as is brought out by (3.11). Even a ${}_{\xi}q_{\eta}$ of 25 0/00 and an R_N' as high as 3.10 will only produce a correction of ${}_{\xi}q_{\eta} \cdot R_N = 0.08$, however. (The value $R_N' = 3.10$ is quoted for Costa Rica for the year 1960 by van de Walle (1967), and is the highest value for R_N' given in his list.) The numerical error caused by the practice mentioned is therefore in most cases certainly of no consequence in comparison with the other sources of error encountered in the process of furnishing values for the reproduction rates.

Nevertheless it has some theoretical interest to note the difference between f_x and f_x', and to be aware of the fact that it represents a slight approximation not to make this distinction.

Table (3.12)

Values of $0.5 \cdot q_x$ for some age between 25 and 30 for females in various populations

Country	Period	Age group	$0.5\ q_x$ in 0/00
Congo (Leopoldville), African population	1950-52	25-29	24.5
Mexico	1940	25-29	22.7
Argentina	1947	25-29	8.3
India	1941-50	30	8.4
Japan	1959	25-29	4.3
Israel	1960	25-29	2.2
Cananda	1955-57	30	0.5
Norway	1951-55	30	0.5
England and Wales	1950-52	30	0.7
Poland	1958	30	0.7

Source: U.N. Demographic Yearbook 1961, Table 25.

4. Maximum likelihood estimation

§ 4.1. (Simplification to constant forces.) Let x be an integral age. For the purpose of estimating the forces of mortality and fertility for ages in the age interval $[x, x+1>$, we shall let the constant parameters μ and ϕ represent the values of the functions $\mu(x+t)$ and $\phi(x+t)$ for $0 \leqq t \leqq 1$. This has practical interest provided $\sup_{0 \leqq t \leqq 1} |\mu(x+t) - \mu|$ and $\sup_{0 \leqq t \leqq 1} |\phi(x+t) - \phi|$ are not too large.

Under these circumstances we have for $0 \leqq t < 1$

$$ {}_tp_x = e^{-\mu t}, \; f_1(x,t) = \phi t, \; P_k(x,t) = \frac{(\phi t)^k}{k!} e^{-(\mu+\phi)t} \quad \text{for } k=0,1,2,\ldots, $$

and $f(x,t) = \frac{\phi}{\mu}(1 - e^{-\mu t})$.

Similarly by (2.10) and after some arithmetic

$$ Q_k(x,t) = \frac{\mu^{k+1}\phi^k}{(\mu+\phi)^{k+1}} \left\{ 1 - \sum_{\nu=0}^{k} \frac{1}{\nu!} \left[\frac{(\mu+\phi)t}{\mu} \right]^{\nu} e^{-(\mu+\phi)t} \right\} \quad \text{for } 0 \leqq t \leqq 1. $$

We introduce $\phi_k = \phi \Pi_k$, and see that $\phi_k \Delta t + o(\Delta t)$ for $\Delta t > 0$ may be interpreted as the probability that a parent of age $x + t$ will have one birth of k children before age $x + t + \Delta t$, and survive to this age. We have

$$ \sum_{k=1}^{K} \phi_k = \phi . \tag{4.1} $$

We shall assume that all ϕ_k and μ are functionally independent.

§ 4.2. (Further definitions.) Assume that a given number n of parents whose lives are independent, are observed as they pass through the ages in the interval $[x, x+1>$. In order to avoid unnecessarily restricting the generality of our results, we shall not assume that we are able to observe all parents throughout the age interval. Let the length of the period of possible observation through ages in $[x, x+1>$ of parent no. j be $z_j \leqq 1$. For this parent, we shall start measuring time at the start of his (or her) personal interval of observation, and shall thus keep track of him (or her) during a time interval $[0, z_j>$.

Perhaps we should stress that $[0, z_j >$ is the interval of possible observation. If the parent dies at time u while under observation, the actual interval of observation will be $[0, u >$, with $u < z_j$.

In our model $z_1, z_2, \ldots, z_n$ will represent given positive numbers. Let B_j be the number of births and let D_j be the number of deaths experienced by parent no. j during $[0, z_j >$. (Note that B_j is not necessarily the number of children born if multiple births are permitted.) Obivously $D_j = 0$ or $D_j = 1$.

If $B_j > 0$, let $T_j(1), T_j(2), \ldots, T_j(B_j)$ be the moments when births occur to parent no. j (in time order). If $D_j = 1$, let $T_j(B_j + 1)$ be the moment when the parent dies. We let

$$U_j = \begin{cases} z_j & \text{if } D_j = 0, \text{ and} \\ T_j(B_j+1) & \text{if } D_j = 1 . \end{cases}$$

Thus $<0, U_j]$ is the actual period of observation for this parent.

If $B_j > 0$, let C_{ji} be the number of children born at time $T_j(i)$, for $1 \leqq i \leqq B_j$, and let $C_j = \sum_{i=1}^{B_j} C_{ji}$ be the total number of children born to parent no. j during $[0, z_j >$. For typographical reasons, we shall sometimes write $\phi(k)$ for ϕ_k.

§ 4.3. <u>(The likelihood.)</u> We shall find the likelihood for parent no. j. We get

$$P'\{ B_j = 0 \text{ and } D_j = 0 \} = e^{-(\mu+\phi)z_j}.$$

Furthermore, for $b_j > 0$,

$$P'\{ B_j = b_j, D_j = 1, \bigcap_{i=1}^{b_j+1} [t_j(i) < T_j(i) < t_j(i) + dt_j(i)] ,$$
$$\text{and } \bigcap_{i=1}^{b_j} (C_{ji} = c_{ji})\}$$

$$= e^{-(\mu+\phi)t_j(1)} . \phi(c_{j1})dt_j(1) . e^{-(\mu+\phi)[t_j(2) - t_j(1)]}\phi(c_{j2})dt_j(2) .$$

$$. \ldots e^{-(\mu+\phi)[t_j(b_j) - t_j(b_j - 1)]}\phi(c_{jb_j})dt_j(b_j) .$$

$$. e^{-(\mu+\phi)[t_j(b_j+1) - t_j(b_j)]}\mu dt_j(b_j+1)$$

$$= e^{-(\mu+\phi)t_j(b_j+1)} .\mu . \prod_{i=1}^{b_j} \phi(c_{ji}) \prod_{i=1}^{b_j+1} dt_j(i).$$

By treating the other cases similarly, we see that the likelihood for parent no. j amounts to

$$F_j = e^{-(\mu+\phi)U_j} \cdot \mu^{D_j} \cdot \prod_{i=1}^{B_j} \phi(C_{ji}) \quad \text{with} \quad \prod_{i=1}^{B_j} \phi(C_{ji}) = 1 \quad \text{if } B_j = 0.$$

Let N_{jk} be the number of the B_j births that have multiplicity k. We then get

$$\ln F_j = -(\mu+\phi)U_j + D_j \ln \mu + \sum_{k=1}^{K} N_{jk} \ln \phi_k$$

where ln denotes the Napierian logaritm. The total log-likelihood for all n independent parents is therefore

$$\ln F = -(\mu+\phi)U + D \ln \mu + \sum_{k=1}^{K} N_k \ln \phi_k, \tag{4.2}$$

where $U = \Sigma U_j$ is the aggregated lifetime observed, $D = \Sigma D_j$ is the total number of deaths observed, and $N_k = \sum_{j=1}^{n} N_{jk}$ is the total number of births observed of multiplicity k. We shall also need the total number of births observed, $B = \sum_j B_j = \sum_k N_k$.

By the properties of Darmois-Koopmans classes of probability distributions, the random vector $(U, D, N_1, \ldots, N_K)$ is minimal sufficient. We note in particular that neither the birth times $T_j(i)$ nor the numbers C_{ji} of children born in the various births enter directly into (4.2).

It is now easily seen that the maximum likelihood estimators for the various parameters are

$$\hat{\mu} = \frac{D}{U}, \hat{\phi}_k = \frac{N_k}{U}, \hat{\phi} = \frac{B}{U}, \text{ and } \hat{\Pi}_k = \frac{N_k}{B}. \tag{4.3}$$

All of these are natural estimators. From (4.3) we derive the maximum likelihood estimators ${}_t\hat{p}_x = e^{-\hat{\mu}t}$, $\hat{f}_1(x,t) = \hat{\phi}t$, $\hat{f}(x,t) = \frac{\hat{\phi}}{\hat{\mu}}(1 - e^{-\hat{\mu}t})$, and $\hat{\Pi} = \sum_{k=1}^{K} k\hat{\Pi}_k = C/B$, where $C = \sum_{ji}\sum C_{ji}$ is the total number of children born during observation. We shall call $\hat{\mu}$ the (age-specific) <u>mortality rate</u> to distinguish it from the corresponding <u>force</u> μ, which it estimates.

§ 4.4. (M.l. estimators for f_x and f'_x.) By (4.3) and the definitions of f_x and f'_x, maximum likelihood estimators for these two parameters are

$$\hat{f}'_x = \hat{\Pi}\hat{f}_1(x,1) = \hat{\Pi}\hat{\phi} = C/U$$

and
$$\hat{f}_x = \hat{\Pi}\hat{f}(x,1) = \hat{\Pi}\hat{\phi}\kappa(\hat{\mu}) = \hat{f}'_x\kappa(\hat{\mu}) \tag{4.4}$$

with
$$\kappa(\mu) = \frac{1}{\mu}(1 - e^{-\mu}). \tag{4.5}$$

Thus $\hat{f}'_x$ is the common age-specific fertility rate. By (4.4) we may regard $\kappa(\hat{\mu})$ as a correction factor which transforms the estimate of partial fertility into an estimate of influenced fertility. We shall call $\hat{f}'_x$ the (age-specific) partial fertility rate, and we shall use the term (age-specific) influenced fertility rate for $\hat{f}_x$.

As $\kappa(\mu) \approx 1 - \frac{1}{2}\mu$, we get $\hat{f}_x \approx \hat{f}'_x(1 - \frac{1}{2}\hat{\mu})$, which is quite similar to (3.10).

§ 4.5. (Probability limit of U/n.) The distribution of U_j may be found as follows: We see that $P\{U_j = z_j\} = P\{D_j = 0\} = e^{-\mu z_j}$. Furthermore, for $0 < t < z_j$, $P\{t < U_j < t + dt\} = e^{-\mu t}.\mu\, dt$. Thus var U_j exists, and

$$EU_j = z_j.e^{-\mu z_j} + \int_0^{z_j} \mu t\, e^{-\mu t}\, dt = \frac{1}{\mu}(1 - e^{-\mu z_j}).$$

Therefore, $\operatorname*{plim}_{n\to\infty} \frac{U}{n} = E\frac{U}{n} = \frac{1}{\mu}(1 - e^{-\mu\bar{z}})$ with $\bar{z}$ defined by

$$e^{-\mu\bar{z}} = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} e^{-\mu z_j}, \tag{4.6}$$

provided the right hand limit of (4.6) exists.

§ 4.6. <u>(Consistency of μ.)</u> Since $P\{D_j = 0\} = e^{-\mu z_j}$ and $P\{D_j = 1\} = 1 - e^{-\mu z_j}$, $ED_j = 1 - e^{-\mu z_j}$, and we get $\operatorname{plim}_{n\to\infty} D/n = E\frac{D}{n} = 1 - e^{-\mu \bar{z}}$. Therefore, $\operatorname{plim}_{n\to\infty} \hat{\mu} = \dfrac{\operatorname{plim}\frac{D}{n}}{\operatorname{plim}\frac{U}{n}} = \mu$. Thus $\hat{\mu}$ is consistent.

§ 4.7. <u>(Consistency of $\hat{\phi}$.)</u> Since $EB_j = f(x,z_j) = \frac{\phi}{\mu}(1 - e^{-\mu z_j})$, we have $\operatorname{plim}_{n\to\infty} \frac{B}{n} = E\frac{B}{n} = \frac{\phi}{\mu}(1 - e^{\mu \bar{z}})$, and $\operatorname{plim}_{n\to\infty} = \dfrac{\operatorname{plim}\frac{B}{n}}{\operatorname{plim}\frac{U}{n}} = \phi$, so $\hat{\phi}$ is consistent.

§ 4.8. <u>(Consistency of $\hat{\Pi}_k$.)</u> Given that $B_j = b$, N_{jk} will be binominally distributed (b,Π_k). Thus $E(N_{jk}|B_j) = B_j.\Pi_k$, and it follows that $\operatorname{plim}_{n\to\infty} \frac{N_k}{n} = E\frac{N_k}{n} = \Pi_k.\ \frac{\phi}{\mu}(1 - e^{-\mu \bar{z}})$. Thus $\operatorname{plim}_{n\to\infty} \hat{\Pi}_k = \dfrac{\operatorname{plim}\frac{N_k}{n}}{\operatorname{plim}\frac{B}{n}} = \Pi_k$, so also each $\hat{\Pi}_k$ is consistent.

§ 4.9. By the results above, the estimators ${}_x\hat{p}_x$, $\hat{f}_1(x,t)$, $\hat{f}(x,t)$, $\hat{f}_x$, $\hat{f}'_x$, $\hat{\Pi}$, and the $\hat{\phi}_k$ are all consisten as $n \to \infty$. Beside this, maximum likelihood estimators in the present kind of model have certain additional optimal characteristics which makes one prefer them to other natural estimators. We shall give one example.

Assume in this paragraph that all z_k have the same value z. By Hoem (1968, § 5.6), $\hat{\phi}_1,\ldots,\hat{\phi}_K$, and $\hat{\mu}$ are asymptotically independent and normally distributed with means $\phi_1,\ldots,\phi_K$, and μ, respectively, and with asymptotic variances

$$\text{as.var } \hat{\mu} = \frac{1}{n} \cdot \frac{\mu^2}{1-e^{-\mu z}} \text{ and as.var } \hat{\phi}_k = \frac{1}{n} \cdot \frac{\mu\phi_k}{1-e^{-\mu z}} \quad \text{for } k = 1,2,\ldots, K.$$

By Sverdrup (1965, Appendix B), these rates are furthermore optimal among Fisher-consistent estimators.

Similar results hold for the other estimators above. Asymptotic properties of $\hat{\phi}$, $\hat{f}_2(x,t)$, $\hat{f}_x$, $\hat{f}'_x$, and the $\hat{\Pi}_k$ will be studied in chapter 7. The asymptotic properties of ${}_t\hat{p}_x$ have been studied by other authors, notably Sverdrup (1961). The properties of $\hat{f}_1(x,t)$ and $\hat{\Pi}$ easily follow from the results for $\hat{\phi}$ and the $\hat{\Pi}_k$ and will not be considered here.

§ 4.10. One of the advantages of an approach where the length z_j of the period of observation of parent no. j may depend on j, is the possibilities it gives for permitting entries into the group of parents studied after the start of the observations. In certain cases it is natural to regard $z_1, z_2,\ldots, z_n$ as values of independent and identically distributed variables. The estimators will still be consistent, and $\hat{\phi}_1,\ldots,\hat{\phi}_K$ and $\hat{\mu}$ will again be asymptotically independent and normally distributed with means $\phi_1,\ldots,\phi_K$, and μ, and with asymptotic variances

$$\text{as.var } \hat{\mu} = \frac{1}{n} \cdot \frac{\mu^2}{1-I}, \text{ as.var } \hat{\phi}_k = \frac{1}{n} \cdot \frac{\mu\phi_k}{1-I} \quad \text{for } k = 1,2,\ldots, K,$$

with $I = \int_0^{z_0} e^{-\mu z}\, dG(z)$, where G is the distribution function of the z_k, and z_0 is the maximal lenght of possible observation, so that $G(z_0) = 1$. The optimality of our rates among Fisher-consistent estimators is preserved at least if G is fully known.

These aspects have been discussed by Hoem (1968) in a wider context. The idea of introducing random z_k is due to Sverdrup (1961, 1965).

§ 4.11. It is also possible to allow for several causes of decrement from the group of parents considered by splitting μ into a corresponding number of forces μ_i. This will leave unchanged such formulae as (4.3) and (4.4) as well as the properties of those estimators. (Note that μ will then be the total force $\Sigma\mu_i$ of decrement from the group.)

Such procedures will make possible more sophisticated methods of fertility analysis than the ones sketched in previous paragraphs, as well as a simultaneous analysis of nuptiality. This is not the subject of the present paper, however.

5. Further notes on estimation

§ 5.1. In the present chapter, it will be assumed that all z_j equal a preassigned value z. A natural estimator for f(x,z) is then the average number of births per parent:

$$\dot{\tilde{f}}(x,z) = \bar{B} = \frac{1}{n}\sum_{j=1}^{n} B_j.$$

A corresponding estimator for f_x if z = 1, is the average number of children per parent:

$$\dot{\tilde{f}}_x = \bar{C} = \frac{1}{n}\sum_{j=1}^{n} C_j = \frac{1}{n}\sum_{k=1}^{K} kN_k,$$

where $C_j = \Sigma\, C_{ji}$ still is the number of children born to parent no. j. By §§ 4.7 and 4.8, both of these estimators are unbiased and consistent. We shall find their variances in §§ 5.3 and 5.4 below.

§ 5.2. If z = 1, a natural estimator for f'_x is similarly the average number of children born to parents who survive age x + 1:

$$\dot{\tilde{f}}'_x = \frac{1}{n-D}\sum_{j=1}^{n} (1 - D_j)\, C_j,$$

provided D < n. In the (improbable) case where D = n, some other formula must be found, e.g. $\tilde{f}'_x = \bar{C}$. Comparison of $\tilde{f}_x$ and $\tilde{f}'_x$ again brings out the distinction between f_x and f'_x.

By previous results, $E(B_j \mid D_j = 0) = f_1(x,1) = \phi$. Moreover, $C_j = \sum_{k=1}^{K} k.N_{jk}$. Thus by § 4.8,

$$E(\tilde{f}'_x \mid D = d) = E(C_j \mid D_j = 0) = \sum_{k=1}^{K} k\, E(N_{jk} \mid D = 0) =$$

$$= \sum_{k=1}^{K} k\, E\{E(N_{jk} \mid B_j, D_j = 0)\} = \sum_{k=1}^{K} k.E(B_j\Pi_k \mid D_j = 0) = \phi \sum_{k=1}^{K} k\Pi_k = \phi\Pi,$$

so $E(\tilde{f}'_x \mid D = d) = f'_x$ for $d = 0,1,\ldots, n-1$. D will be binominally distributed (n,p) with $p = e^{-\mu}$, so $P(D = n) = e^{-\mu n}$. Thus

$$E\tilde{f}'_x = f'_x\,.(1 - e^{-\mu n}) + E(\tilde{f}'_x \mid D = n)\,.e^{-\mu n},$$

and $\tilde{f}'_x$ is generally at least asymptotically unbiased as $n \to \infty$.

Since $E(1 - D_j)\, C_j = E(C_j \mid D_j = 0)\, P(D_j = 0) = \phi\Pi\,.\, e^{-\mu}$, we get

$$\operatorname{plim} \tilde{f}'_x = \frac{\operatorname{plim} \frac{1}{n} \sum_{j=1}^{n} (1 - D_j)C_j}{\operatorname{plim} (1 - \frac{D}{n})} = \phi\Pi = f'_x,$$

so $\tilde{f}'_x$ will also be consistent. We may find its variance, but the mathematical formula is ugly and probably of little use. One should prefer $\hat{f}'_x$ to $\tilde{f}'_x$ in any case.

§ 5.3. Since $\operatorname{var} \tilde{f}(x,t) = \frac{1}{n} \operatorname{var} B_j$, we need a formula for $\operatorname{var} B_j$. We are then primarily interested in $g(x,t) = \sum_{k\geq 1} k^2\, R_k(x,t) = E\, B_j^2$.

If $g_1(x,t) = \sum_{k\geq 1} k^2\, P_k(x,t)/{}_tp_x$, and $g_2(x,t) = \sum_{k\geq 1} k^2\, Q_k(x,t)/{}_tq_x$, then $g(x,t) = {}_tp_x\, g_1(x,t) + {}_tq_x \cdot g_2(x,t)$. The formulae of § 4.1 give

$$g_1(x,t) = \sum_{k=1}^{\infty} k^2 \frac{(\phi t)^k}{k!} e^{-\phi t} = (\phi t)^2 + \phi t.$$

The combination of this result with (2.10) gives

$$g_2(x,t) \cdot {}_tq_x = \int_0^t g_1(x,\tau) \cdot {}_\tau p_x \cdot \mu d\tau.$$

$$= \frac{\phi}{\mu}\left(2\frac{\phi}{\mu} + 1\right) - e^{-\mu t}\left\{(\phi t)^2 + \left(2\frac{\phi}{\mu} + 1\right)\left(\phi t + \frac{\phi}{\mu}\right)\right\}.$$

Thus $g(x,t) = (1 - e^{-\mu t})\frac{\phi}{\mu}\left(2\frac{\phi}{\mu} + 1\right) - 2e^{-\mu t}\frac{\phi}{\mu}\phi t = EB_j^2,$

and
$$\operatorname{var} B_j = g(x,t) - \left(\frac{\phi}{\mu}\right)^2 (1 - e^{-\mu t})^2. \tag{5.1}$$

It is possible to show that, when all $z_j = z$,

$$\operatorname{var} \tilde{f}(x,t) \geq \text{as.var}\, \hat{f}(x,t), \tag{5.2}$$

so $\hat{f}(x,t)$ would be preferred. (Cfr. Sverdrup (1965, Appendix B).) $\tilde{f}_x$ is Fisher-consistent.) The combination of (5.1) and (5.2) gives an upper bound for as.var $\hat{f}(x,t)$. A formula for this asymptotic variance will be given as (7.4).

§ 5.4. To find var $\tilde{f}_x = \frac{1}{n}$ var C_j, we return to § 2.7 and note that

$$\operatorname{var}(C_{ji} | B_j \geq i) = \sum_{k=1}^{K} k^2 \Pi_k - \Pi^2.$$

Thus if $\tau^2 = \sum_{k=1}^{K} k^2 \Pi_k$, $\operatorname{var}(C_j | B_j = b) = b(\tau^2 - \Pi^2)$. In § 2.7 we similarly found $E(C_j | B_j = b) = b\Pi$. Thus

$$\operatorname{var} C_j = E \operatorname{var}(C_j | B_j) + \operatorname{var} E(C_j | B_j) = (\tau^2 - \Pi^2) EB_j + \Pi^2 \operatorname{var} B_j$$

or
$$\operatorname{var} C_j = \tau^2 \frac{\phi}{\mu}(1 - e^{-\mu}) + \left(\frac{\Pi\phi}{\mu}\right)^2 (1 - 2\mu \cdot e^{-\mu} - e^{-2\mu}).$$

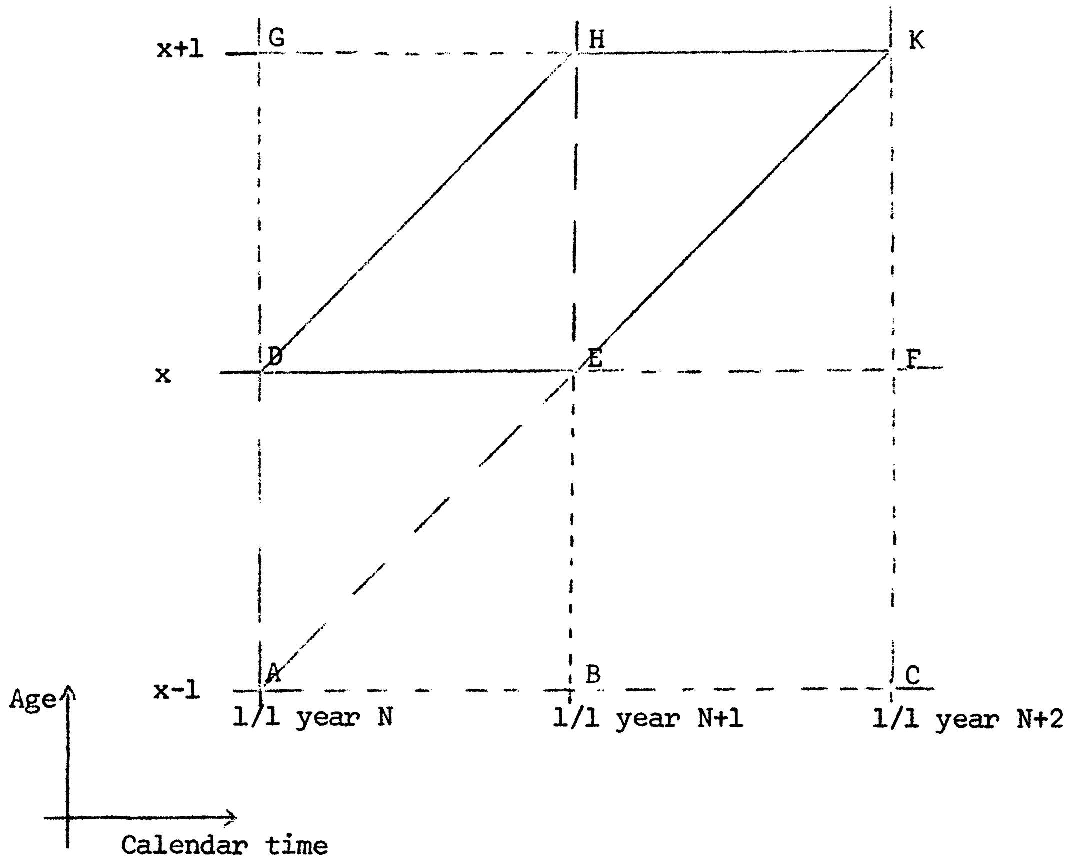
6. Some applications to estimation in demography

§ 6.1. The application of the results of the previous chapters to demographic estimation is fairly straightforward. We shall give three examples to indicate how this may be done. Adaption to other situations can be made ad hoc.

To fix our ideas, we shall concentrate on a situation where it is desired to investigate female fertility with respect to (live) children of both sexes in a population. Let x be an integer. We define

Figure 6.1

Section of the Lexis diagram.



$L_x^{(N)}$ as the number of women of age x in the population on January 1 of the year N,

$M_x^{(N)}$ as the number of women who experience their x-th birthday in the population in the year N,

$C_x^{(N)+}$ as the number of live children born in the year N by mothers who experience their x-th birthday in that year and who have the birth no earlier than this birthday, and

$C_x^{(N)-}$ as the number of live children born in the year N by x-year old mothers who have the birth before their own birthday in that year[1].

Thus $L_x^{(N)}$ is the number of female lifelines that cross the line DG in figure 6.1, and $M_x^{(N)}$ is the number of female lifelines that cross DE. Similarly, $C_x^{(N)+}$ is the number of children born "in the triangle DEH", and $C_x^{(N)-}$ is the corresponding number of children in Δ DHG.

Quantities $D_x^{(N)+}$ and $D_x^{(N)-}$ will be defined quite similarly to $C_x^{(N)+}$ and $C_x^{(N)-}$, respectively, and will signify numbers of female deaths. Similarly $U_x^{(N)+}$ and $U_x^{(N)-}$ will signify aggregated female lifetimes.

Thus e.g. $U_x^{(N)+}$ is the total lenghts of the female lifelines in Δ DEH in figure 6.1.

We shall consider the estimation of fertility in the parallelograms DEKH, AEHD, and DEHG. As a side result we will simultaneously get mortality estimates.

§ 6.2. <u>(Age year method.)</u> It may be desired to analyse fertility in the parallelogram DEKH. Let $f_x^{(N)}$ be the quantity corresponding to the f_x of (3.5) for a woman whose lifeline has points in the parallelogram, let $f_x^{(N)'}$ be the quantity corresponding to the f_x' of (3.7), and let $\mu_x^{(N)}$ be her force of mortality (regarded as a constant parameter throughout the parallelogram). By (4.3) and § 4.4 one would estimate these three quantities by

$$\hat{\mu}_x^{(N)} = \frac{D_x^{(N)+} + D_x^{(N+1)-}}{U_x^{(N)+} + U_x^{(N+1)-}},$$

$$\hat{f}_x^{(N)'} = \frac{C_x^{(N)+} + C_x^{(N+1)-}}{U_x^{(N)+} + U_x^{(N+1)-}}, \qquad (6.1)$$

and $$\hat{f}_x^{(N)} = \hat{f}_x^{(N)'} \cdot \kappa(\hat{\mu}_x^{(N)}),$$

1) At this birthday, these mothers will become x+1 years old.

If the data are not available in a form which gives the relatively detailed information required in these estimation formulae, various approximation techniques can be used. For example

$$\tfrac{1}{2}\left[M_x^{(N)} + M_{x+1}^{(N+1)}\right] \text{ may be chosen as an approximation to } U_x^{(N)+} + U_x^{(N+1)-}.$$

If not even the $M_x^{(N)}$ are available, their values again may be approximated by similar techniques. One may e.g. choose $\tfrac{1}{2}\left[L_{x-1}^{(N)} + L_x^{(N+1)}\right]$ as an approximation to $M_x^{(N)}$. In this case, $U_x^{(N)+} + U_x^{(N+1)-}$ will be approximated by

$$\frac{1}{4}L_{x-1}^{(N)} + \frac{1}{2}L_x^{(N+1)} + \frac{1}{4}L_{x+1}^{(N+2)}.$$

The quantity corresponding to the $\tilde{f}_x$ of § 5.1 will be

$$\dot{\tilde{f}}_x^{(N)} = \frac{C_x^{(N)+} + C_x^{(N+1)-}}{M_x^{(N)}}. \qquad (6.2)$$

§ 6.3. <u>(Calendar year method.)</u> Let ${}^{*}f_x^{(N)}$, ${}^{*}f_x^{(N)'}$ and ${}^{*}\mu_x^{(N)}$ be the quantities corresponding to the f_x of (3.5), the f_x' of (3.7), and the force of mortality, respectively, for a woman whose lifeline has points in AEHD. Then

$$\begin{aligned} {}^{*}\hat{\mu}_x^{(N)} &= \frac{D_{x-1}^{(N)-} + D_x^{(N)+}}{U_{x-1}^{(N)-} + U_x^{(N)+}}, \\ {}^{*}\hat{f}_x^{(N)'} &= \frac{C_{x-1}^{(N)-} + C_x^{(N)+}}{U_{x-1}^{(N)-} + U_x^{(N)+}}, \end{aligned} \qquad (6.3)$$

and $${}^{*}\hat{f}_x^{(N)} = {}^{*}\hat{f}_x^{(N)'} \cdot \kappa({}^{*}\mu_x^{(N)}).$$

If necessary, $U_{x-1}^{(N)-} + U_x^{(N)+}$ may be approximated by $M_x^{(N)}$ or by $\frac{1}{2}\left[L_{x-1}^{(N)} + L_x^{(N+1)}\right]$. The relation corresponding to (6.2) is

$$x\tilde{f}_x^{(N)} = \frac{C_{x-1}^{(N)-} + C_x^{(N)+}}{L_{x-1}^{(N)}}. \tag{6.4}$$

§ 6.4. Let ${}^{\dagger}f_x^{(N)}$, ${}^{\dagger}f_x^{(N)'}$ and ${}^{\dagger}\mu_x^{(N)}$ be the quantities corresponding to the f_x of (3.5), the f_x' of (3.7), and the force of mortality, respectively, for a woman whose lifeline has points in DEHG. Then

$$
\begin{aligned}
{}^{\dagger}\hat{\mu}_x^{(N)} &= \frac{D_x^{(N)+} + D_x^{(N)-}}{U_x^{(N)+} + U_x^{(N)-}}, \\
{}^{\dagger}\hat{f}_x^{(N)'} &= \frac{C_x^{(N)+} + C_x^{(N)-}}{U_x^{(N)+} + U_x^{(N)-}}, \qquad (6.5)
\end{aligned}
$$

and $\quad {}^{\dagger}\hat{f}_x^{(N)} = {}^{\dagger}\hat{f}_x^{(N)'} \cdot \kappa({}^{\dagger}\hat{\mu}_x^{(N)})$.

An approximation to $U_x^{(N)+} + U_x^{(N)-}$ suggested by Sverdrup (1961) is

$$\frac{1}{6} L_{x-1}^{(N)} + \frac{1}{3} L_x^{(N)} + \frac{1}{3} L_x^{(N+1)} + \frac{1}{6} L_{x+1}^{(N+1)}.$$

It is an essential feature of the present situation that not all z_j are equal so there is no formula corresponding to (6.2) and (6.4).

§ 6.5. It appears that fertility rates are commonly estimated by formulae like (6.1), (6.3), and (6.5), generally with some approximation included. This amounts to estimating partial fertility. (Cfr. § 3.2).

7. Asymptotic properties of maximum likelihood estimators

§ 7.1. Let $\Theta = (\Theta_1, \ldots, \Theta_s)'$ be any parameter vector, let $\hat{\Theta} = (\hat{\Theta}_1, \ldots, \hat{\Theta}_s)'$ be a maximum likelihood estimator for Θ, and assume that $\sqrt{n}\,(\hat{\Theta} - \Theta)$ is asymptotically multinormal with mean 0 and some covariance matrix Σ. (This will generally though not always be the case for maximum likelihood estimators.) We introduce a new parameter vector $\alpha = (\alpha_1, \ldots, \alpha_s)'$ by a one-to-one transformation ψ of the form $\Theta_k = \psi_k(\alpha_1, \ldots, \alpha_s)$ for $k = 1, 2, \ldots, s$. The differential quotients $\frac{\partial}{\partial \alpha_j} \psi_k(\alpha_1, \ldots, \alpha_s)$ are assumed to exist and be continuous. We introduce

$$J = \begin{Bmatrix} \frac{\partial \psi_1}{\partial \alpha_1}, & \cdots & , \frac{\partial \psi_1}{\partial \alpha_s} \\ \cdot & & \cdot \quad \cdot \\ \frac{\partial \psi_s}{\partial \alpha_1}, & \cdots & , \frac{\partial \psi_s}{\partial \alpha_s} \end{Bmatrix},$$

and let $\hat{\alpha}$ be the maximum likelihood estimator for α. It may then be shown that $\sqrt{n}\,(\hat{\alpha} - \alpha)$ is asymptotically multinormal with mean 0 and some covariance matrix $\Gamma = (\gamma_{ij})$. Hoem (1968, § 6.1) shows that

$$\gamma^{ij} = \sum_{\nu=1}^{s} \sum_{k=1}^{s} \sigma^{\nu k} \cdot \frac{\partial \psi_\nu}{\partial \alpha_i} \cdot \frac{\partial \psi_k}{\partial \alpha_j}, \quad \text{for all } i \text{ and } j, \tag{7.1}$$

where $\Gamma^{-1} = (\gamma^{ij})$ and $\Sigma^{-1} = (\sigma^{ij})$. This formula may also be written $\Gamma^{-1} = J'\Sigma^{-1}J$.

§ 7.2. We studied asymptotic properties of the estimator $\hat{\Theta} = (\hat{\phi}_1, \ldots, \hat{\phi}_K, \hat{\mu})'$ in §§ 4.9 and 4.10, and found that $\sqrt{n}\,(\hat{\Theta} - \Theta)$ will be asymptotically normal with mean 0 and asymptotic covariance matrix $\beta\Delta$, where

$$\beta = \begin{cases} \mu/(1 - e^{-\mu z}) \text{ under the assumptions of § 4.9, and} \\ \mu/\{1 - \int_0^{z_0} e^{-\mu z}\, dG(z)\} \text{ under the assumptions of § 4.10,} \end{cases}$$

and $\Delta = \operatorname{diag}(\phi_1, \phi_2, \ldots, \phi_K, \mu)$.

In the present chapter it does not matter which one of these two sets of assumptions apply. Our results hold in either case. Formula (7.1) will now reduce to

$$\gamma^{ij} = \frac{1}{\beta} \sum_{k=1}^{K+1} \frac{\partial \psi_k}{\partial \alpha_i} \quad \frac{\partial \psi_k}{\partial \alpha_j} \Big/ \phi_k, \tag{7.2}$$

where $\phi_{K+1} = \mu$.

§ 7.3. <u>(Properties of $\hat{\mu}$, $\hat{\phi}$, and the $\hat{\Pi}_k$.)</u> We have defined K to be the maximal multiplicity possible of a birth. The special case of K = 1 is rather trivial and will not be considered here. In the rest of this chapter, we shall let $K \geqq 2$, so that it is possible to get at least twins in any birth.

Since $\sum_{k=1}^{K} \Pi_k = 1$, there are K - 1 "free" parameters Π_k. In the present paragraph we shall study asymptotic properties of $\hat{\Pi}_1, \ldots, \hat{\Pi}_{K-1}$ along with $\hat{\phi}$ and $\hat{\mu}$.

Since $\phi = \sum_{k=1}^{K} \phi_k$, $\Pi_k = \phi_k / \phi$, we have the situation of § 7.1 with $\alpha = (\Pi_1, \ldots, \Pi_{K-1}, \phi, \mu)$ and $\psi_k(\alpha) = \phi\Pi_k$ for $k = 1,2,\ldots. K-1$, $\psi_K(\alpha) = \phi(1 - \sum_{k=1}^{K-1} \Pi_k)$, $\psi_{K+1}(\alpha) = \mu$. Thus $\sqrt{n}(\hat{\alpha} - \alpha)$ will be asymptotically multinormal with mean 0 and some covariance matrix Λ which we intend to find. We have

$$J = \begin{Bmatrix} \phi, & 0, & 0, & \ldots & , 0, & \Pi_1, & 0 \\ 0, & \phi, & 0, & \ldots & , 0, & \Pi_2, & 0 \\ \cdot & \cdot & & \cdot\ \cdot & \cdot & & \cdot \\ 0, & 0, & 0, & \ldots & , \phi, & \Pi_{K-1}, & 0 \\ -\phi, & -\phi, & -\phi, & \ldots & , -\phi, & 1 - \sum_{k=1}^{K-1} \Pi_k, & 0 \\ 0, & 0, & 0, & \ldots & , 0, & 0, & 1 \end{Bmatrix}$$

Applying (7.2), we get:

$$\lambda^{ij} = \begin{cases} \frac{\phi}{\beta}\left(\frac{\delta_{ij}}{\Pi_i} + \frac{1}{\Pi_K}\right) & \text{for } 1 \leqq i < K,\ 1 \leqq j < K, \\ \frac{1}{\beta\phi} & \text{for } i = j = K, \\ \frac{1}{\beta\mu} & \text{for } i = j = K+1, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{7.3}$$

Here δ_{ij} is a common Kronecker delta. If Λ_1^{-1} is the (K-1) x (K-1) matrix of the λ^{ij} of the first line of (7.3) and $\Lambda_2^{-1} = \text{diag}\left(\frac{1}{\beta\phi}, \frac{1}{\beta\mu}\right)$, then

$$\Lambda^{-1} = \begin{Bmatrix} \Lambda_1^{-1}, & 0 \\ 0, & \Lambda_2^{-1} \end{Bmatrix} \quad \text{and } \Lambda = \begin{Bmatrix} \Lambda_1, & 0 \\ 0, & \Lambda_2 \end{Bmatrix}.$$

Thus the three random variables $(\hat{\Pi}_1, \ldots, \hat{\Pi}_{K-1})'$, $\hat{\phi}$, and $\hat{\mu}$ are asymptotically independent, and

$$\text{as.var } \hat{\phi} = \phi\beta/n, \text{ while as.var } \hat{\mu} = \mu\beta/n \text{ as in § 7.2.}$$

Inverting Λ_1^{-1} we also find after some calculation:

$$\text{as.var } \hat{\Pi}_k = \Pi_k(1 - \Pi_k)\,\beta/(n\phi), \text{ and}$$

$$\text{as.cov } (\hat{\Pi}_j, \Pi_k) = -\Pi_j \Pi_k \,\beta/(n\phi).$$

§ 7.4. Now let $\theta = (\Pi_1, \ldots, \Pi_{K-1}, \phi, \mu)'$. To study the properties of $\hat{f}(x,t)$ we introduce $\alpha_K = \Pi_k$ for $k = 1,2,\ldots, K-1$, $\alpha_K = f(x,t) = \frac{\phi}{\mu}(1 - e^{-\mu t})$, and $\alpha_{K+1} = \mu$. In the terminology of § 7.1, $\psi_k(\alpha_1, \ldots, \alpha_{K+1}) = \alpha_k$ for $k = 1,\ldots, K-1$, and $K+1$, while $\psi_K(\alpha_1, \ldots, \alpha_{K+1}) = \alpha_K \alpha_{K+1}(1 - e^{-t\alpha_{K+1}})^{-1}$. We let $a = \mu(1-e^{-\mu t})^{-1}$, $b = \frac{\phi}{\mu}(1 - at^{-\mu t})$ and

$$A = \begin{Bmatrix} a, & b \\ 0, & 1 \end{Bmatrix}, \quad \text{and get} \quad J = \begin{Bmatrix} I, & 0 \\ 0, & A \end{Bmatrix},$$

where I is a $(K - 1) \times (K - 1)$ identity matrix. Thus

$$\sqrt{n}\,(\hat{\Pi}_1 - \Pi_1, \ldots, \hat{\Pi}_{K-1} - \Pi_{K-1}, \hat{f}(x,t) - f(x,t), \hat{\mu} - \mu)$$

is asymptotically multinormal with mean 0 and a covariance matrix $\mathcal{A} = (J' \Lambda^{-1} J)^{-1}$, where Λ is given as in § 7.3. Since $J' \Lambda^{-1} J = \begin{Bmatrix} \Lambda_1^{-1}, & 0 \\ 0, & A'\Lambda_2^{-1}A \end{Bmatrix}$, we see that $(\hat{\Pi}_1, \ldots, \hat{\Pi}_{K-1})$ and $(\hat{f}(x,t), \hat{\mu})$ are asymptotically independent, $(\hat{\Pi}_1, \ldots, \hat{\Pi}_{K-1})$ has the asymptotic covariance matrix $\frac{1}{n}\Lambda_1$ as in § 7.3, and $(\hat{f}(x,t), \hat{\mu})$ has the asymptotic covariance matrix $\frac{1}{n}(A' \Lambda_2^{-1} A)^{-1}$. Some trivial arithmetic gives

$$\text{as.var } \hat{f}(x,t) = (\phi + \mu b^2)\,\beta/(na^2), \text{ and} \tag{7.4}$$

$$\text{as.cov } (\hat{f}(x,t), \hat{\mu}) = -\mu\beta b/(na),$$

while as.var $\hat{\mu} = \mu\beta/n$ as before.

§ 7.5. With $\theta = (\Pi_1, \ldots, \Pi_{K-1}, \phi, \mu)'$, we now introduce $\alpha_k = \Pi_k$ for $k = 1, 2, \ldots, K-1$, $\alpha_K = f_x'$, and $\alpha_{K+1} = f_x$. We shall conclude this chapter by studying the asymptotic properties of $\hat{f}_x'$ and $\hat{f}_x$.

In the terminology of § 7.1, $\psi_k(\alpha_1, \ldots, \alpha_{K+1}) = \alpha_k$ for $k = 1, 2, \ldots, K-1$, $\psi_K(\alpha_1, \ldots, \alpha_{K+1}) = \alpha_K / \{K - \sum_{k=1}^{K-1} (K-k)\, \alpha_k\}$, and $\psi_{K+1}(\alpha_1, \ldots, \alpha_{K+1}) = \kappa^{-1}(\alpha_{K+1}/\alpha_K)$. Letting

$$D = -\frac{\phi}{\Pi} \begin{Bmatrix} K-1, & K-2, & \ldots, & 1 \\ 0, & 0, & \ldots, & 0 \end{Bmatrix} \text{ and}$$

$$G = \begin{Bmatrix} \dfrac{1}{\Pi}, & 0 \\ -\dfrac{\kappa(\mu)}{\kappa'(\mu) f_x'}, & \dfrac{1}{\kappa'(\mu) f_x'} \end{Bmatrix},$$

we get $J = \begin{Bmatrix} I, & 0 \\ D, & G \end{Bmatrix}$, where I is a $(K-1) \times (K-1)$ identity matrix and $\kappa'(\mu) = \dfrac{d}{d\mu}\kappa(\mu) = \dfrac{1}{\mu}\left[e^{-\mu} - \kappa(\mu)\right]$. Defining $H = J' \Lambda^{-1} J$, we have

$$H = \begin{Bmatrix} \Lambda_1^{-1} + D' \Lambda_2^{-1} D, & D' \Lambda_2^{-1} G \\ G' \Lambda_2^{-1} D, & G' \Lambda_2^{-1} G \end{Bmatrix}$$

Here

$$D' \Lambda_2^{-1} D = \frac{\phi}{\beta \Pi^2} \begin{Bmatrix} (K-1)^2, & (K-1)(K-2), & (K-1)(K-3), & \ldots, & K-1 \\ (K-2)(K-1), & (K-2)^2, & (K-2)(K-3), & \ldots, & K-2 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ K-1, & (K-2), & (K-3), & \ldots, & 1 \end{Bmatrix},$$

$$G' \Lambda_2^{-1} D = \frac{1}{\beta \phi \Pi} D, \text{ and } G' \Lambda_2^{-1} G = \frac{c}{\mu \beta} \begin{Bmatrix} d, & -\kappa(\mu) \\ -\kappa(\mu), & 1 \end{Bmatrix},$$

with $c = 1/\left[\kappa'(\mu)\, f_x'\right]^2$ and $d = \kappa^2(\mu) + \phi\mu\left[\kappa'(\mu)\right]^2$ (7.5)

The vector $(\hat{\Pi}_1, \ldots, \hat{\Pi}_{K-1}, \hat{f}_x', \hat{f}_x)$ will be asymptotically multinormal with mean $(\Pi_1, \ldots, \Pi_{K-1}, f_x', f_x)$ and covariance matrix $\frac{1}{n} H^{-1}$.

We shall not invert H for general K, but shall be content with the case K = 2. We then account for twin births but rule out triplets and births with even higher multiplicity. In human populations, such births are rare in any case.

When K = 2, $\Lambda_1^{-1} = \frac{\phi}{\beta}\left(\frac{1}{\Pi_1} + \frac{1}{\Pi_2}\right) = -\frac{\phi}{\beta\Pi_1\Pi_2}$, $D'\Lambda_2^{-1}D = \frac{\phi}{\beta\Pi^2}$, and $D'\Lambda_2^{-1}G = -\frac{1}{\beta\Pi^2}(1,0)$. Thus

$$H = \frac{1}{\beta}\begin{Bmatrix} \phi g, & -\Pi^{-2}, & 0 \\ -\Pi^{-2}, & cd/\mu, & -\kappa(\mu)\, c/\mu \\ 0, & -\kappa(\mu)\, c/\mu, & c/\mu \end{Bmatrix}$$

with
$$g = \frac{1}{\Pi_1\Pi_2} + \frac{1}{\Pi^2}. \tag{7.6}$$

Thus

$$H^{-1} = \beta\Pi_1\Pi_2 \cdot \begin{Bmatrix} 1/\phi, & 1, & \kappa(\mu) \\ 1, & \Pi^2\phi g, & \Pi^2\phi g\kappa(\mu) \\ \kappa(\mu), & \Pi^2\phi g\,\kappa(\mu), & \phi\Pi^2\left[g\kappa^2(\mu) + \phi\mu(\kappa'(\mu))^2/(\Pi_1\Pi_2)\right] \end{Bmatrix}$$

For K = 2 we therefore get

$$\text{as.var } \hat{f}_x' = \frac{1}{n} \cdot f_x'\, \beta\Pi(1 + \Pi_1\Pi_2/\Pi^2),$$

$$\text{as.var } \hat{f}_x = \frac{1}{n}\beta\Pi\left[f_x(1 + \Pi_1\Pi_2/\Pi^2)\kappa(\mu) + f_x'\,\phi\mu(\kappa'(\mu))^2\right]$$

and an asymptotic covariance between $\hat{f}_x'$ and $\hat{f}_x$ equal to $\frac{1}{n}\beta\Pi f_x(1 + \Pi_1\Pi_2/\Pi^2)$.

8. Appendix

§ 8.1. In § 2.4 we considered a parent known to be alive at age x, and studied conditional probabilities of events occurring in the interval $[x, x+t]$, given that the parent would survive to age x + t. We asserted that such probabilities and the corresponding mean values could be evaluated by simply letting $\mu(x + \tau) = 0$ for $0 \leq \tau \leq t$.

§ 8.2. Although intuitively plausible, such an assertion is not obvious and it needs a proof, all the more so as a seemingly corresponding assertion is not correct in ordinary multiple-decrement theory. Specifically the "partial probability" $q_{x,\nu}$ of § 3.2 may not be interpreted as the probability that a person of age x will die before age x + 1 of cause ν, given that he will not die of one of the other causes before age x + 1. The latter, conditional probability in fact quals $q_x^{(\nu)} / \{q_x^{(\nu)} + p_x\}$, which is generally quite different from $q_{x,\nu}$.

A proof of the assertion of § 8.1 is given below.

§ 8.3. The model of chapter 2 is a time-continuous age-dependent Markov chain with a double infinity of states. A parent is studied over the age interval $[x, x + t]$, say. In the following let $0 \leq \tau \leq t$. The parent will have a number k of births at ages in the interval $[x, x + \tau]$. If the parent is still alive at age $x + \tau$, we shall say that he then belongs to state (k,1). Otherwise he belongs to state (k,2). Designating his state at age y by S(y), we define

$$P_{ij}(y, u) = P'\{S(y + u) = j \mid S(y) = i, S(x) = (0,1)\} \quad \text{for } y \geq x.$$

The force of transition from state i to state j at age y is defined as

$$\mu_{ij}(y) = \lim_{u \downarrow 0} P_{ij}(y, u) / u \quad \text{for } i \neq j.$$

Then $P_k(x,t) = P_{(0,1),(k,1)}(x,t)$, $Q_k(x,t) = P_{(0,1)(k,2)}(x,t)$, $\mu(x) = \mu_{(k,1),(k,2)}(x)$, $\phi(x) = \mu_{(k,1),\ (k+1,1)}(x)$ for all $k \geq 0$, and $\mu_{ij} = 0$ otherwise. Probabilities conditional on the assumption that the parent will still be alive at age x + t, are introduced by

$$A_{ij}(y, u) = P'\{ S(y + u) = (j, 1) \mid S(x) = (0,1), S(y) = (i,1), B \},$$

for $x < y \leq y + u \leq x + t$, where B is the event that the parent actually survives age $x + t$. For $j \geq i$ we get

$$P'\{ S(y) = (i, 1), S(y + u) = (j, 1), B \mid S(x) = (0,1) \} =$$

$$= P_i(x, y - x) P_{j-i}(y, u) \cdot {}_{x+t-(y+u)}p_{y+u}$$

and $P'\{ S(y) = (i,1), B \mid S(x) = (0,1) \} = P_i(x, y - x) \cdot {}_{x+t-y}p_y$.
Thus

$$A_{ij}(y, u) = P_{j-i}(y, u) \cdot {}_{x+t-(y+u)}p_{y+u} / {}_{x+t-y}p_y.$$

We therefore get

$$\lim_{u \downarrow 0} A_{ij}(y, u) / u = \delta_{j,i+1} \cdot \phi(y) \quad \text{for } j \neq i.$$

This means, however, that the $A_{ij}(y, u)$ may be regarded as the transition probabilities of a simple age-dependent birth process with intensity $\phi(y)$, as was our assertion.

9. Acknowledgement

I am grateful to Mr. Bjørn L. Tønnesen, who has proof-read the manuscript of this paper.

10. References

Consael, R. and A. Lamens (1962): "Méthodes déterministes et probabilistes de l'analyse démographique". Biométrie-Praximétrie, 3, 83-112.

Hoem, Jan M. (1968): "Applications of Time-Continuous Markov Chain Estimation Techniques to Demography". Under publication.

Joshi, D.D. (1954): "Les processus stochastiques en démographie". Publ. Inst. Stat. Univ. Paris, 3, 149-176.

Kendall, M.G. and W.R. Buckland (1957): "A Dictionary of Statistical Terms". Oliver and Boyd.

Khintchine, A.Y. (1960): "Mathematical Methods in the Theory of Queuing". Charles Griffin & Co. Ltd.

Mathiessen, P.C. (1967): "Teoretisk Demografi". Københavns Universitets Fond til Tilvejebringelse af Læremidler.

Sverdrup, Erling (1961): "Statistiske metoder ved dødelighetsundersøkelser". Mimeographed. Inst. of Math., Univ. of Oslo.

Sverdrup, Erling (1965): "Estimates and Test Procedures in connection with Stochastic Models for Deaths, Recoveries and Transfers between different States of Health". Skand. Aktuarietidskrift 1965, 184-211.

United Nations: "Demographic Yearbook 1961".

van de Walle, Etienne (Associate Editor for Statistics) (1967): "Measures of Fertility, Mortality and Reproduction of Female Populations". Population Index, 33 (2), 287-293.