J. Math. Biol. 48, 105–118 (2004)
Digital Object Identifier (DOI):
10.1007/s00285-003-0231-9

Mathematical Biology

J.D.H. Smith

A macroscopic approach to demography

Received: 16 May 2001 / Revised version: 4 May 2003 /
Published online: 20 August 2003 –

Abstract. A *canonical/lognormal* model for human demography is established, specifying the net maternity function and the age distribution for mothers of new-borns using a single macroscopic parameter vector of dimension five. The age distribution of mothers is canonical, while the net maternity function normalizes to a lognormal density. Comparison of an actual population with the model serves to identify anomalies in the population which may be indicative of phase transitions or influences from levels outside the demographic. Tracking the time development of the parameter vector may be used to predict the future state of a population, or to interpolate for data missing from the record. In accordance with classical theoretical considerations of Backman, Prigogine, *et al.*, it emerges that the logarithm of a mother's age is the most fundamental time variable for demographic purposes.

1. Introduction

One of the recurring difficulties of demographic analysis has been the complexity of the data required to give an adequate specification of a particular population. Typically, this complexity is not less than the order of the number of age groups into which the population is classified. For example, the number of non-zero entries in a projection matrix is about twice the number of age classes [10, p.41]. In other words, something like a score of "microscopic" coefficients are required to describe the state of the population. The current paper is intended to initiate a macroscopic approach to demography, based on a *canonical/lognormal model*, specifying a given population by a five-dimensional parameter vector from which detailed information such as net maternity function values may be generated. The development of the population over time may then be tracked kinetically merely by following the trajectory of this five-dimensional vector. Under certain smoothness conditions with a wide range of validity, one may interpolate for times at which the historical data are incomplete, or readily predict the future development of the population.

It is important to note that no assumptions concerning Lotka stability are required by the canonical/lognormal model, in contrast with many of the previously available demographic techniques. Indeed, one of the five parameters of the model is the dimensionless *perturbation* which explicitly measures the deviation from Lotka

J.D.H. Smith: Dept. of Mathematics, Iowa State University, Ames, IA 50011, U.S.A.
e-mail: `jdhsmith@math.iastate.edu`; *URL*: http://www.math.iastate.edu/jdhsmith/

Key words or phrases: Net maternity function – Lotka stability – Canonical ensemble – Lognormal distribution – Malthusian parameter – Maximum entropy principle – Organic time – Thermodynamic time

stability. Thus Lotka stability is merely a special case of the canonical/lognormal model, much as stationarity is a special case of Lotka stability.

From one point of view, the macroscopic approach may be regarded as an exercise in data compression. From another, it mimics the thermodynamic concept of an ideal gas, in which detailed information about the positions and momenta of the 10^{27} gas molecules in a large chamber may be summarized by the pressure, volume, and temperature. Although the data compression to be afforded by the techniques of the present paper is by no means comparable with that achieved by the ideal gas concept — after all, human beings are much more complex than gas molecules — the analogy is still useful in helping one appreciate what will and will not be done here.

The first point concerns the distinction between intensive and extensive variables. Intensive variables are those, such as pressure and temperature, that do not depend on the size of the population sample. Extensive variables, such as volume, are those that do depend on the size of the sample. In the demographic context, the current paper focuses on intensive variables such as net maternity function values, not on extensive variables such as the total female population.

The second point concerns the distinction between model and reality. The ideal gas is a simple model. Real gases, such as unsuperheated steam, do not behave exactly like an ideal gas. In similar fashion, real populations will not behave exactly as predicted by the simple models given in this paper. Sometimes, as for the 1985 Malaysian population described in Figure 1, the agreement will be close. At other times, e.g. for the 1985 U.S. population described in Figure 3, there will be marked discrepancies. These discrepancies are useful indicators of the presence of additional factors not built in to the model. The discrepancy between the behaviors of an ideal gas and of unsuperheated steam is partly due to condensation (phase transition) of the latter. The discrepancy between the behavior of a model population with the macroscopic parameters of the U.S. in 1985, and that of the actual population, is due to peculiarities of the actual population, for example the deferment of motherhood by women pursuing a career. Beyond the data compression aspect, one of the prime applications of the current model is to locate special features of a given population, features that may not be immediately apparent from the mass of raw data with which one is initially confronted. From a complex systems perspective, these features may be identified as influences on a population's demographic level exerted by other levels, such as the organismic level from below or the economic level from above.

2. Plan of the paper

The present paper is concerned with macroscopic specification of two intensive functions of age a (years) in a human female population: the net maternity function $\varphi(a)$ and the probability density function $q(a)$ for the age of the mother of a randomly chosen new-born. There are five macroscopic parameters that determine these functions: a constant R_{-1} (per annum) normalizing $\varphi(a)$ into a density (4.5); a Malthusian parameter r (per annum); a dimensionless parameter s measuring the perturbation from Lotka stability; a logarithmic generation time t (years); and a dimensionless standard deviation u (4.11). In terms of these parameters, the net maternity function has a *lognormal* density

$$\varphi(a) = \frac{R_{-1}}{u\sqrt{2\pi}} \exp\left\{ -\frac{1}{2u^2}\left(\log\frac{a}{t}\right)^2\right\}$$

[cf. (4.13)], while the probability density function $q(a)$ is determined by (the normalization $\int_0^\infty q(a)da = 1$ and) the proportionality

$$q(a) \propto \exp\left\{ -ra - \frac{1+s}{2u^2}\left(\log\frac{a}{t}\right)^2\right\}.$$

The density $q(a)$ is *canonical*. One may refer to the population model under study as the *canonical/lognormal model*. Canonical distributions are discussed in Section 3, using a continuous age variable for convenience. Fuller details, explicitly using discrete age classes, are available in earlier papers [15] [16]. The lognormal distribution for the net maternity function, and the resulting canonical distribution, are presented in Section 4. For further details about lognormal distributions, including various biological applications and potential generating mechanisms, a useful reference is [5].

The latter half of the paper examines the validity of the canonical/lognormal model, and discusses its use as a tool for the analysis of actual populations. The discussion divides naturally into two aspects: demographic statics (Section 5) and demographic kinetics (Section 6). Demographic statics is concerned with the analysis of a population at a particular moment. Here, the key question is the descriptive power of the canonical/lognormal model. Generally, the fit is observed to be very good, especially for populations not subject to marked stress. Indeed, significant divergence of the actual functions from the models is a useful indicator of the presence of such stresses, e.g. the deferment of motherhood in the 1985 U.S. population. A major feature of the model is the way it provides a basis for comparison entirely within the static framework, not needing data values from other time points. For example, comparison of the 1985 U.S. data with the model may replace comparison of the 1985 data with their 1980 counterparts.

Demographic kinetics is concerned with the evolution of populations over time. The canonical/lognormal model enables one to focus on the trajectory of the five-dimensional vector (R_{-1}, r, s, t, u) of parameters, much as the ideal gas model enables one to study the behavior of a gas in terms of pressure, temperature, and volume changes. Under normal circumstances (e.g. the absence of shock waves in the ideal gas), the macroscopic variables change smoothly with time. One may use this assumption of smoothness in the time evolution of the parameter vector to predict the future state of a population. Section 6 presents an example of this, examining the (peninsular) Malaysian population over the 1970–1985 time period. As a test, the 1970–1980 data are used to predict the state in 1985. Comparison of the prediction with the actual state shows very close agreement. Such predictions are essentially extrapolations. In similar vein, one could use assumed smoothness of parameter change to interpolate states into gaps in the sequence of available data.

The paper concludes with a brief discussion of future research problems raised by the present work. The two main issues are:

(1) Development of a demographic dynamics to govern the kinetics of the parameter vector (R_{-1}, r, s, t, u);
(2) Finding an intrinsic biological foundation for the logarithmic time scale of the net maternity function.

Problem (1) seems unapproachable at present, although it is hoped that long-term tracking of certain well-documented populations may start to reveal some of the forces involved. Problem (2) suggests some intriguing connections with classical ideas of Backman [3] and Prigogine [13], ideas that have started to reappear in recent work such as [1].

3. The canonical distribution

Consider the experiment of choosing a new-born female at random, and determining the age A of her mother. Suppose that the probability density function for this random variable A is $p(a)$, so that the probability of a random new-born's mother having age lying in the range $t_1 < A \leq t_2$ is

$$\int_{t_1}^{t_2} p(a)da. \tag{3.1}$$

Since $p(a)$ is a probability density function, it is normalized by

$$1 = \int_0^{\infty} p(a)da. \tag{3.2}$$

Define the *generation time*

$$T = \int_0^{\infty} ap(a)da \tag{3.3}$$

as the expected age of a random new-born's mother. Define the *(logarithmic) maternity*

$$M = -\int_0^{\infty} p(a)\log\varphi(a)da \tag{3.4}$$

as the expected value of the negated logarithm of the *net maternity function* $\varphi(a)$ [8, p.100]. Note that the quotient M/T is the *reproductive potential* [7]. Suppose that the numerical values of the generation time and logarithmic maternity (or generation time and reproductive potential) are determined, but that one has no further information about the probability density function $p(a)$. In this case, the appropriate model for the density function $p(a)$ is the *canonical density function* $q(a)$, namely the (unique) density function that maximizes the *entropy* [14, §20] or expected value

$$H = -\int_0^{\infty} p(a)\log p(a)da \tag{3.5}$$

of the negated logarithm of the density function, subject to the constraints (3.2), (3.3), and (3.4). Introduce a Lagrange multiplier α corresponding to the constraint (3.2). Introduce a Lagrange multiplier r corresponding to the generation time constraint (3.3). The quantity r is known as the *Malthusian parameter*; its units are inverse time (e.g. per year). Introduce a Lagrange multiplier $(1 + s)$ corresponding to the maternity constraint (3.4). The dimensionless quantity s is known as the *perturbation*. The Euler equation for the constrained maximization problem is

$$-\log p(a) - 1 - \alpha - ra + (1 + s)\log \varphi(a) = 0. \tag{3.6}$$

Thus its solution, the canonical density function $q(a)$, is given by the equation

$$\log q(a) = -\log Z(r, s) - ra + (1 + s)\log \varphi(a) \tag{3.7}$$

as

$$q(a) = Z(r, s)^{-1} e^{-ra} \varphi(a)^{1+s}. \tag{3.8}$$

Here the *partition function* or Zustandsumme

$$Z(r, s) = \int_0^\infty e^{-ra} \varphi(a)^{1+s}\, da \tag{3.9}$$

is given by

$$\log Z(r, s) = 1 + \alpha \tag{3.10}$$

in terms of the Lagrange multiplier α. Denote the *Lotka* (*stable*) *growth rate* by r_1 [10, p.41]. The Lotka characteristic equation [8, §6.5] [9, §5.1] [11, p.65] then takes the form

$$Z(r_1, 0) = 1 \tag{3.11}$$

according to (3.9). From (3.8), it follows that the Lotka-stable density

$$q^0(a) = e^{-r_1 a} \varphi(a) \tag{3.12}$$

[6, (2.8)] is the special case of the canonical density obtained by setting the perturbation s to be zero and the Malthusian parameter r to be equal to the Lotka growth rate r_1. (This state of affairs may be summarized by referring to Lotka stability as the *unperturbed* case $r = r_1$, $s = 0$ of canonicity, much as stationarity is the special case of Lotka stability corresponding to $r_1 = 0$.) Now take the negated expected value of each side of (3.7) with respect to the canonical probability distribution. One obtains

$$H = \log Z(r, s) + rT + (1 + s)M. \tag{3.13}$$

This equation holds exactly for the particular r and s values determined by the Lagrange multipliers. The numerical values of these parameters are small (of the order of $1/T$). Now although $\log Z(r, s)$ is a strictly convex function (cf. [2, Cor. 2.2]), its graph differs little from its tangent planes in the demographically significant region of small r and s values. To within demographic accuracy, one may thus assume that one has a linear function

$$\log Z(r, s) = (H - M) - rT - sM \tag{3.14}$$

for small r and s, namely $O(1/T)$. Setting $r = r_1$ and $s = 0$ into (3.14), and using (3.11), one thus extends the relationship

$$r_1 = \frac{H - M}{T} \tag{3.15}$$

from the Lotka-stable case [6, (2.16)] to the general canonical case. Note that (3.15) yields r_1 more directly than does solution of the characteristic equation.

4. The lognormal distribution

Setting $s = 0$ into equation (3.9), one obtains

$$Z(r, 0) = \sum_{k=0}^{\infty} \frac{(-r)^k}{k!} R_k \tag{4.1}$$

with

$$R_k = \int_0^{\infty} a^k \varphi(a) da \tag{4.2}$$

for natural numbers k [8, §6.1] [9, §5.2]. In particular, one has the *net reproduction rate*

$$R_0 = Z(0, 0) \tag{4.3}$$

[8, p.102]. Although the net maternity function $\varphi(a)$ is non-negative, it does not directly yield a probability density function on the set of ages unless stationarity holds. In this latter case $r_1 = 0$, and then (3.11) and (4.3) together imply $R_0 = 1$. In the general case one might renormalize, considering

$$\varphi(a)/R_0 \tag{4.4}$$

as a density [9, §5.2]. Attempts have been made to fit (4.4) with normal and other readily handled density functions on the set of ages [8, Ch. 6]. However, as Keyfitz observes [8, p. 168], the attempts have not generally been satisfactory.

The essence of the current approach is to consider an appropriate renormalization of the net maternity function as a probability density function for the logarithm of the age, not for the age itself. The *renormalization constant* must be $\int_{a=0}^{a=\infty} \varphi(a) d(\log a)$ or

$$R_{-1} = \int_0^{\infty} a^{-1} \varphi(a) da \tag{4.5}$$

[extending the notation (4.2)]. One then obtains

$$\varphi(a)/R_{-1} \tag{4.6}$$

as a probability density function. Under this density, the probability for the age random variable X to lie in a range $t_1 < X \le t_2$ is given as

$$\int_{a=t_1}^{a=t_2} \frac{\varphi(a)da}{R_{-1}a}. \tag{4.7}$$

One may also set $Y = \log X,\ y = \log a$ and use the equivalent expression

$$\int_{y=\log t_1}^{y=\log t_2} R_{-1}^{-1}\varphi(e^y)dy. \tag{4.8}$$

The expected value of the logarithm of the age under the density (4.6) will be written in the form

$$\log t = \int_0^\infty \frac{\varphi(a)\log a}{R_{-1}a} da \tag{4.9}$$

for a characteristic age t called the *logarithmic generation time*. Generally, one has

$$t < T. \tag{4.10}$$

(The mnemonic "little t less than big T" is sometimes helpful.) The variance of the logarithm of the age under the density (3.6) will be written in the form

$$u^2 = \int_0^\infty \frac{\varphi(a)[\log(a/t)]^2}{R_{-1}a} da \tag{4.11}$$

with a non-negative, dimensionless parameter u called the (*standard*) *deviation*. Thus the random variable $Y - \log t$ has mean 0 and standard deviation u. Assuming that one has no further information about the distribution of Y, one concludes that $Y = \log X$ is normally distributed with mean $\log t$ and variance u^2 [14, p. 629]. Equivalently, the age random variable X has the two-parameter lognormal distribution

$$\Lambda(\log t, u). \tag{4.12}$$

In other words [5, p.2], the net maternity function is given as

$$\varphi(a) = \frac{R_{-1}}{u\sqrt{2\pi}} \exp\left\{ -\frac{1}{2u^2}\left(\log\frac{a}{t}\right)^2\right\} \tag{4.13}$$

for $a > 0$. Note its dependence on the three parameters: the renormalization constant R_{-1}, the logarithmic generation time t, and the standard deviation u. Making further use of the Malthusian parameter r and perturbation s, one may insert (4.13) into (3.8) to obtain

$$q(a) = \exp\left\{ -ra - \frac{1+s}{2u^2}\left(\log\frac{a}{t}\right)^2\right\} \Big/ \int_0^\infty \exp\left\{ -ra - \frac{1+s}{2u^2}\left(\log\frac{a}{t}\right)^2\right\}da \tag{4.14}$$

or

$$q(a) = Z(r,s)^{-1}\Big(\frac{R_{-1}}{u\sqrt{2\pi}}\Big)^{1+s} \exp\Big\{ -ra - \frac{1+s}{2u^2}\Big(\log\frac{a}{t}\Big)^2\Big\} \tag{4.15}$$

with

$$Z(r,s) = \Big(\frac{R_{-1}}{u\sqrt{2\pi}}\Big)^{1+s} \int_0^\infty \exp\Big\{ -ra - \frac{1+s}{2u^2}\Big(\log\frac{a}{t}\Big)^2\Big\} da. \tag{4.16}$$

Together, (4.13) and (4.14) specify $\varphi(a)$ and $q(a)$ entirely in terms of the vector

$$(R_{-1}, r, s, t, u) \tag{4.17}$$

of parameters. More succinctly, one may describe the shapes of $\varphi(a)$ and $q(a)$ by the proportionalities

$$\varphi(a) \propto \exp\Big\{ -\frac{1}{2u^2}\Big(\log\frac{a}{t}\Big)^2\Big\} \tag{4.18}$$

and

$$q(a) \propto \exp\Big\{ -ra - \frac{1+s}{2u^2}\Big(\log\frac{a}{t}\Big)^2\Big\}. \tag{4.19}$$

The canonical/lognormal model may then be summarized by (4.5), (4.18) and (4.19).

5. Demographic statics

This section illustrates the application of the canonical/lognormal model to what one might call “demographic statics” — the analysis of a certain population at one given time. In [16, §5], the canonical distribution was used for such an analysis of the population of Malaysian females in 1970. The canonical distribution (3.8) overestimated the births to mothers in the 25 to 29 year old age class by an order of 10%. This was the cohort of mothers born under the occupation of (then) Malaya during World War II. Thus the discrepancy between the true record of births and that estimated by the canonical model pointed to an anomaly in the population due to historical effects not built in to the model.

The first of the current analyses takes the same population 15 years later: (peninsular) Malaysian females in 1985. The data are presented in Figure 1, based on [10, pp.384-5]. The graphs displayed are piecewise-linear rather than smooth, due to the recording of the data in 5 year classes. The upper panel graphs are normalized to display total births (of both sexes) for the various 5-year age classes, corresponding to the “births by age of mother” tabulations in [10]. The lower panel graphs are normalized to display the per quinquennium net maternity function, in accordance with the tabulations in [10]. The solid graphs, with data points marked by plus signs, give the true births by age (class) and the true net maternity function. For example, the plus signs over age 27.5 years record the actual number of

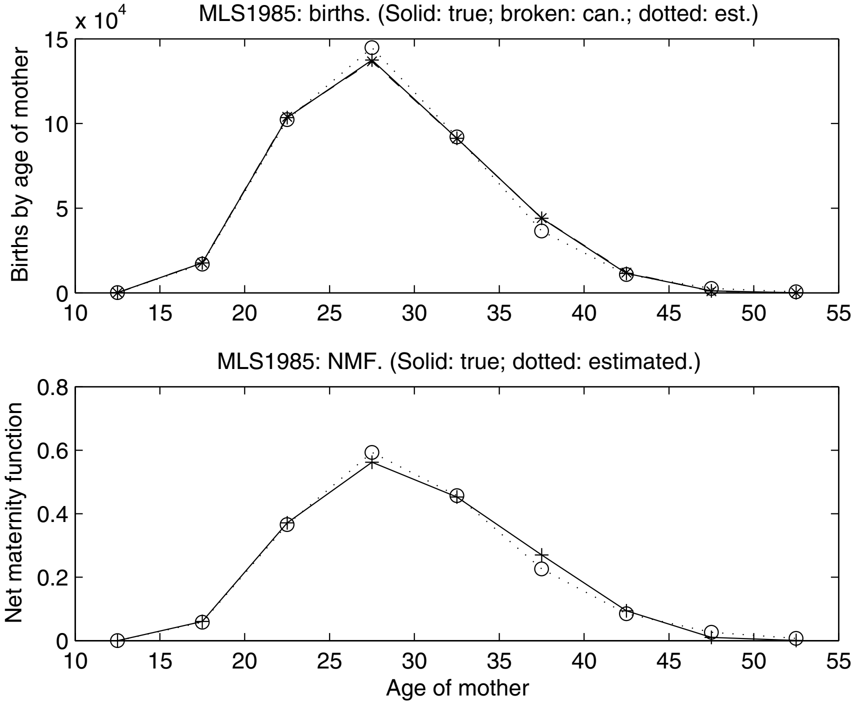

Fig. 1. Malaysia 1985.

13.747×10^4 births for mothers in the 25–29 year age class, and the corresponding net maternity function value of 0.5620 per quinquennium. On the graph of births, the canonical distribution (3.8) using the true net maternity function is indicated by broken lines, with crosses marking data points. The fit is so exact that the discrepancy between the true and canonical graphs is barely discernible. The agreement suggests that the population has returned to equilibrium after the disturbance of World War II. Since the perturbation value s at 0.0733 is significantly different from zero, however, one can not characterize this equilibrium as Lotka stability. The dotted graph in the lower panel, with circles marking data points, displays the net maternity function estimate given by the lognormal model (4.13). Finally, the dotted graph in the upper panel gives the estimated births by age using the fully parametrized canonical distribution (4.14). The fit to the net maternity function is quite good, being slightly too peaked. These mild distortions are reflected in the fit of the parametrized canonical distribution.

The second figure (Figure 2) displays the corresponding graphs for Cypriot females in 1980, based on [10, pp.358-9]. Possibly due to the relative smallness of the sampled population, or to unreliability of the data because of the partition of the island, the true births by mother's age and the net maternity function have the curious feature of not being unimodular: there are 2 births recorded for the 45–49 year age class, but 4 for the 50–54 year olds. Although not apparent from the figure,

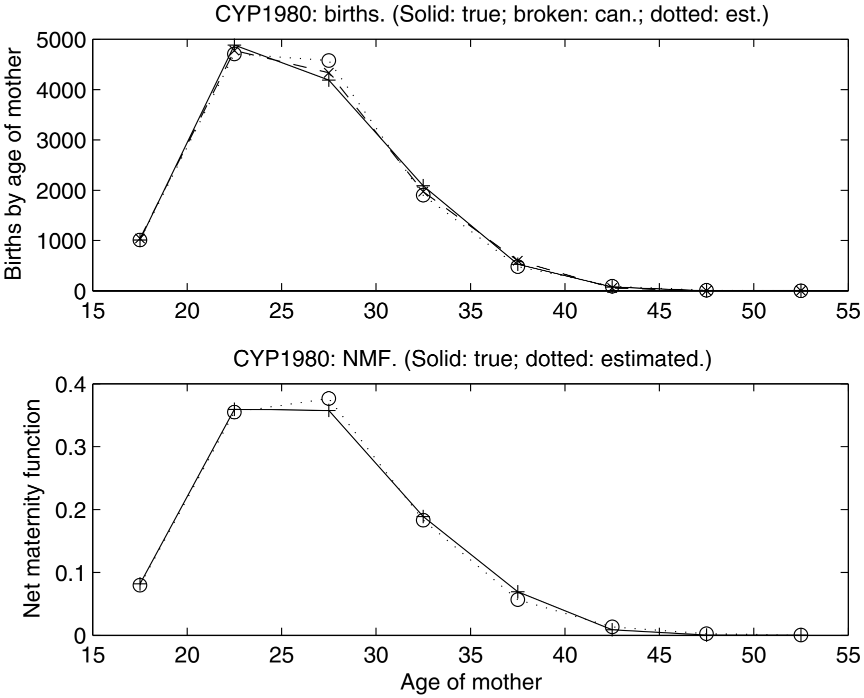

Fig. 2. Cyprus 1980.

the canonical model (3.8) using the true net maternity function also displays this lack of unimodularity, estimating 1.3 births to the 45–49 year olds and 2.1 to the 50–54 year olds. These anomalies are smoothed out in the parametrized curves, again shown dotted in each panel. As before, the fits are quite good. This time, however, the parametrized canonical distribution for the births by mother's age is not as peaked as the true curve.

The final figure of the section (Figure 3) displays the corresponding data for United States females in 1985, based on [10, pp. 348-9]. The (broken) canonical graph using true net maternity function values gives a very close tracking of the true births by mother's age. Unlike the case of 1970 Malaysia, this graph does not indicate any anomalies. However, the (dotted) graphs of parameter-based estimates show two distinct discrepancies: a dearth of actual births in the 20–24 year age class, balanced by an excess of actual births in the 30–34 year age class. These discrepancies appear to indicate a significant number of women postponing childbirth for up to ten years, presumably in order to concentrate on a career. Note that the current techniques are indicating these discrepancies on a purely static basis, taking only the data for 1985. Against this background, it is interesting to observe that the graphs of "Changes in Fertility 1950–1985" of [10, p.349] present a similar picture. Comparing the true net maternity functions for 1980 and 1985, that for 1985 is low for women in their early twenties, and high for women in their early thirties.

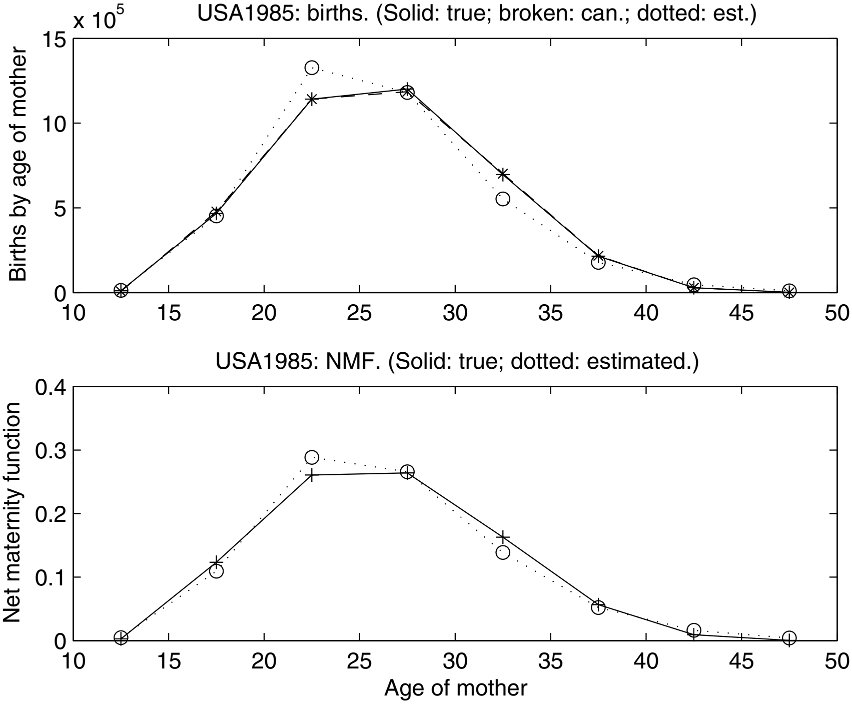

Fig. 3. U.S.A. 1985.

6. Demographic kinetics

The previous section showed how the canonical/lognormal model may be used for a static analysis of a population, working with the data taken from a single point in time. Static canonical/lognormal analysis of U.S. females in 1985 provided information (on the postponement of births by women in their twenties) that was otherwise apparent only by comparison with the state of the population at other times. This section shows how the canonical/lognormal model may be used in "demographic kinetics" – the study of the development of a population over time. In essence, the canonical/lognormal model reduces complete knowledge of the net maternity function and age distribution of mothers to knowledge of the five parameters displayed in the vector (4.17). Under normal circumstances (e.g. in the absence of "phase transitions"), one may expect the vector (4.17) of parameters to evolve smoothly over time. This assumption becomes the basis of a method for predicting the future development of a population. The method will be illustrated by the example of the (peninsular) Malaysian female population over the 1970–1985 time period [10, pp.380-5]. The first four rows of Table 1 display the true values of the components of (4.17) for five-year intervals of the period. Table 2 displays the corresponding true values of various dependent quantities. [Note that the Lotka growth rate r_1 is computed using (3.15).] Now suppose that in 1980 one wished

Table 1. Independent parameters for Malaysian females.

	R_{-1} (/yr)	r (/yr)	s	t (yrs)	u
True 1970	0.0781	0.0369	−0.0718	27.09	0.2361
True 1975	0.0602	0.0415	−0.0225	27.04	0.2331
True 1980	0.0642	0.0400	0.0256	27.24	0.2220
True 1985	0.0642	0.0343	0.0733	27.81	0.2150
Est. 1985	0.0630	0.0322	0.0725	27.69	0.2022
Error	−2%	−6%	−1%	−1%	−6%

Table 2. Dependent quantities for Malaysian females.

	$\log Z$	r_1 (/yr)	M	T (yrs)	H
True 1970	−0.2164	0.0268	0.9166	27.83	1.6622
True 1975	−0.5069	0.0223	1.0062	27.55	1.6216
True 1980	−0.5859	0.0198	1.0193	27.75	1.5687
True 1985	−0.4710	0.0203	0.9719	28.48	1.5498
Est. 1985	−0.4245	0.0196	0.9283	28.30	1.4819
Error	10%	−3%	−4%	−1%	−4%

to predict the condition of the population in 1985. For each of the five parameters of vector (4.17), one would know their three values at the respective times 1970, 1975, 1980. Fitting a quadratic polynomial function of time to these three values, one could then evaluate the polynomial at 1985 to obtain a prediction of the true parameter value in 1985. These five estimates of the 1985 values are displayed in the fifth row of Table 1. The errors range up to 6%.

The estimated parameter values may then be used in formulas (4.13) and (4.14) to give predictions for the 1985 net maternity function and distribution of births by age of mother. The predictions are quite accurate, especially for the distribution of births by age of mother. From these predicted functions, one may derive predictions for the various dependent quantities. These predictions are recorded in the fifth row of Table 2. For example, the error in the predicted Lotka growth rate is −3%. (A direct quadratic fit to the r_1 values gives a 5% underestimate.)

7. Discussion

The work presented above raises two open problems that may serve as the focus for future research: biological modelling of the lognormality of the net maternity function, and development of a "demographic dynamics" to accompany the "demographic kinetics" of Section 6. Taking the latter problem first, recall that the kinetics merely traces the time development of the fundamental parameter vector (4.17). Dynamics would identify forces driving this development, and quantify their effect on the parameters. The problem may be hard to solve, because of the influence of factors from outside the demographic level. These factors may come from lower levels such as that of the organism, e.g. through diseases, or from higher levels such as the economy, e.g. through the development of pension schemes that reduce the incentive to raise a large family. Some attempts directed at developing demographic dynamics have been undertaken by Demetrius, for example (see [7] for a summary), taking the population entropy as the basic parameter. Although this work gives an interesting analysis of the convergence to the stable age distribution in the Lotka model (cf. [17]), it does not appear to be applicable in more general contexts. Consider the 1970–1985 Malaysian population as detailed in Table 2. Note that M and r_1 are positive throughout. On the other hand, the quotient H/T (cf. [7, [9]]) decreases steadily: its four values in time order are 0.0597, 0.0589, 0.0565, 0.0544. This contradicts the prediction of [7, Table 1] that positivity of r_1 and M/T (negativity of $\Phi = -M/T$ as in [7, [9]]) should imply an increase of the normalized entropy H/T. Of course, the Malaysian population is far from stability, since r differs from r_1 throughout, and s differs from 0 except for some point in 1978. The methodology of the current paper suggests that the population entropy is best treated as a dependent variable, the components of (4.17) being the basic parameters.

The lognormality of the net maternity function should certainly be the subject of further investigation. The derivation in Section 3 proceeded from the reasonable epistemological assumption that the only information one had was the mean and variance of the logarithm of the age. It would be interesting to obtain an intrinsic biological model for this distribution. There are two indications that the fundamental variable should be the logarithm of the age, rather than the age itself. The first is that the logarithm of the age functions as an "organic time" in Backman's sense [3, (6)] [13, XIII§3.3] or a "thermodynamic time" in Prigogine's sense [13, (13.21)], cf. [4, p.231]. Indeed, (4.8) with (4.13) may be viewed as an instance of Backman's general growth formula [3, (5)]. (Strictly speaking, this formula applies to a particular organism. For current purposes, one could take the "organism" here to be the cohort of mothers born in a certain time period, together with their daughters considered as "inert tissue" of the organism in the sense of [12, §2.3].) The second indication comes from analyzing dimensions. The net maternity function $\varphi(a)$ has inverse time as its dimension. If one wishes to normalize it to a dimensionless function of age, one should use $\varphi(a)/R_{-1}$ rather than $\varphi(a)/R_0$, thereby making the logarithm of the age into the basic variable.

References

1. Andresen, B., Shiner, J.S., Uhliger, D.E.: Allometric scaling and maximum efficiency in physiological eigen time. Proc. Nat. Acad. Sci. **99**, 5822–5824 (2002)
2. Athreya, K.B., Smith, J.D.H.: Canonical distributions and phase transitions. Discuss. Math. Prob. Stat. **20**, 167–176 (2000)
3. Backman, G.: Lebensdauer und Entwicklung. Arch. für Entwicklungsmechanik **140**, 90–123 (1940)
4. von Bertelanffy, L.: General System Theory. Braziller, New York, 1968
5. Crow, E.L., Shimizu, K.: Lognormal Distributions. Marcel Dekker, New York, 1988

6. Demetrius, L.: Growth rate, population entropy and perturbation theory. Math. Biosci. **93**, 159–180 (1989)
7. Demetrius, L.: Directionality principles in thermodynamics and evolution. Proc. Nat. Acad. Sci. 94, 3491–3498 (1997)
8. Keyfitz, N.: Introduction to the Mathematics of Population. Addison-Wesley, Reading, MA, 1968
9. Keyfitz, N.: Applied Mathematical Demography. 2nd ed., Springer, New York, 1985
10. Keyfitz, N., Flieger, W.: World Population Growth and Aging. University of Chicago, Chicago, 1990
11. Lotka, A.J.: Théorie Analytique des Associations Biologiques II: Analyse Démographique avec Application Particulière á l'Espèce Humaine. Hermann, Paris, 1939
12. Mosimann, J.E. Campbell, G.: Applications in biology: simple growth models. In: "Lognormal Distributions" E.L. Crow and K. Shimizu, (eds.), Marcel Dekker, New York, 1988, pp. 287–302
13. Prigogine, I.: Etude Thermodynamique des Phénomènes Irréversibles. Dunod, Paris, 1947
14. Shannon, C.E.: A mathematical theory of communication. Bell System Tech. J. **27**, 623–656 (1948)
15. Smith, J.D.H.: Competition and the canonical ensemble. Math. Biosci. **133**, 69–83 (1996)
16. Smith, J.D.H.: Demography and the canonical ensemble. Math. Biosci. **153**, 151–161 (1998)
17. Tuljapurkar, S.D.: Why use population entropy? It determines the rate of convergence. J. Math. Biol. **13**, 325–337 (1982)