CUI MR-T-79-006   c.3

THE ANALYSIS OF AGE DISTRIBUTIONS

Daniel Goodman

Scripps Institution of Oceanography
La Jolla, California 92093

November, 1979

(Goodman; R/F-56)

ABSTRACT

The stable age distribution of a life table reflects
survivorship discounted by population growth rate. Graphical
techniques may be employed to estimate the growth rate of
the population from partial information on a population's
age distribution and mortality rates for a few age classes.
When a time varying harvest or some other perturbation
causes departures from the stable age structure, stable
population theory no longer applies, but a kind of convergence
can still be demonstrated. We develop a formula which may be
used to estimate vital rates by comparison of an observed
and a predicted age distribution when the history of the
absolute (not per capita) harvests is known.

ACKNOWLEDGEMENTS

KEY WORDS

age distribution, census, harvest, life table, stable age distribution, transient age distribution

INTRODUCTION

A population's age distribution reflects a history
of the population birth rates corrected for the history of
age specific survival rates. The presence of an individual
aged x at time t arises through birth at time t-x coupled
with survival at least until age x. Relative abundances
of individuals in age classes x and y must, therefore,
indicate relationships between the birth rates at times
t-x and t-y and the probabilities of survival till ages
x and y at those times. Accordingly, appropriate analysis
of an age distribution should shed some light on other
aspects of the population's dynamics--notably birth rates
and death rates.

This fact is of significance in wildlife or fisheries
management, and in certain areas of field ecology, because
relative measures, such as age distributions are more
readily estimated from samples than are absolute measures,
such as direct estimates of population size.

Some of the internal relationships involving age
distributions and various population parameters and age
specific parameters in a stable population are well known.
For example, chapters 2 and 4 of Coale's (1972) monograph
develop the mathematics in a very clear fashion. That
examination of a population structure may yield practical

information about a population is an established idea in fisheries and wildlife work. Relative abundances of young animals are loosely interpretted as indices of the growth or decline of populations. The relationship between certain mortality rates and population composition is exploited in the ingenious survey-removal method of population assessment (Chapman and Murphy, 1965). And, of course, many of the same phenomena are at the heart of various sorts of widely used sequential analyses of population structures, termed virtual population analysis, in fisheries management (Gulland, 1965; Pope, 1972).

Many existing applications of the analysis of age structures employ notation which obscures the relationship to the formulations of mathematical demography. Indeed, some of the methods introduce inaccuracies through neglect of one or the other of population growth rate, natural mortality rates, or transients in the population age distribution. The purposes of this paper are to review the problem from a general demographic perspective, to demonstrate the unappreciated power of the methods in some otherwise conventional applications to stable population analysis, and to break new ground in the analysis of age structures under a fluctuating realized life table.

Two examples--framed in the context of commonly
encountered practical problems--will be explored. One is
the estimation of population growth rate from sparse knowledge
of survival rates and age structure in a population in
stable age distribution. The second is the reconstruction
of some missing elements of the life table from knowledge
of the current age distribution and a history of variable
harvests.

I. ASSESSING POPULATION STATUS IN THE ABSENCE OF FECUNDITY DATA

The two obvious means of assessing the growth rate
of a population are by direct measurements of population
size over a suitable time interval, or by calculation from
the fecundity and survivorship schedule of the applicable
life table. The two methods, in principle will yield the
same answer, provided the population's age distribution
has converged to a distribution reasonably close to the
stable age distribution associated with the life table.
Often enough, neither method can be satisfactorily applied
to populations of interest, notably, stocks of pelagic
animals.

The usual method of estimating population size in
a fishery is by calculation from data on the catch rate
per unit of fishing effort. The results, of course, are
confounded by changes in fishing technology, changes in
the economics of the fishery, and changes in the availability
of fish to the fishery (as distinct from changes in
numbers). The method is reviewed by Gulland (1969).

A second method of estimating population size in
a population that is inaccessible to direct census is by
calculation from the rate of recapture of a known number
of marked individuals relative to the capture of unmarked

ones. This method, which is more usually applied to
terrestrial animals, assumes uniform mixing of the marked
individuals into the population at large. Furthermore,
the results are confounded by the rate of mortality
during the period intervening between release and
recapture. Indeed, where the marked individuals can be
aged, this method serves better as a means of calculating
mortality rates over the respective age classes, than as
a means of estimating population. The method is reviewed
by Seber (1973).

Growth rate may be calculated from the life table
according to Lotka's equation

$$\int_0^\infty e^{-rx} \ell_{(x)} m_{(x)} dx = 1 \quad , \quad (1)$$

where r is the per capita population growth rate, m(x) is
the age specific schedule of fecundity (usually female
births per female, the male segment of the population
is modeled differently), and l(x) is the rate of survival
from birth to age x. The classic development of this
formula is reviewed by Lotka (1939). In practise, a
discrete time analog is employed in place of the continuous
functions, for real data. When reproduction is seasonal,
a discrete time model may, in any case, be more appropriate.

A thorough, modern treatment of demographic methods is given in Keyfitz (1969).

It is only rarely that birth rate data might readily be gathered for a pelagic population. Usually, measures of pregnancy rates or egg burden are the most that can be expected . The survivorships of the younger age classes are almost always beyond reach of feasible measurement. So much so, in fact, that it is common to fill in this missing element in a life table by assuming that r is some particular value (usually 0), and then calculating the rate of immature survivorship that is consistent with this value, given the rest of the life table. Little of practical significance is gained by this exercise, other than perhaps a test of internal consistency or plausibility, for the growth rate is the more telling statistic.

Interestingly, the internal relationships connecting age specific parameters to population parameters permit calculation of r from an incomplete life table, provided at least some elements of the stable age distribution are known. The purpose of this section is to develop the calculation for the realistic case where neither birth rates nor early survival rates are available, and where only certain age classes are accesible to reliable sampling.

THE STABLE AGE DISTRIBUTION

The stable age distribution is an expression of the survivorship schedule, distorted exponentially with age according to the rate of increase of the population. That is, the frequency of age class x is proportional to the simple product of the survivorship to age x and the population's birth rate at the time x units ago. If the population has been growing exponentially at rate r, the birth rate at time (t-x) must be e^{-rx} times the present population birth rate. Upon normalization, therefore, we obtain as the formula for $c^*(x)$, the frequency of age class x,

$$c^*(x) = \frac{e^{-rx} \, l(x)}{\int_0^n e^{-ry} l(y) \, dy} \qquad (2)$$

The denominator in this expression, incidentally, must be the inverse of the mean percapita birth rate in the population. For this reason, it is not necessary to measure births directly.

From (2) we see that the growth rate, r, is determined uniquely by the survivorship schedule, l(x), and the stable age distribution $c^*(x)$. Given l(x) and $c^*(x)$, r may be computed iteratively from equation (2). Alternatively, as

noted by Bourgeois-Pichat (1957), r may be obtained from
the coefficient of the exponential regression of
$(c(x)/l(x))$ against age. Here we will show that even less
information than the complete $c(x)$ and $l(x)$ schedules
will suffice.

MORTALITY RATES AND A PARTIALLY SPECIFIED SURVIVORSHIP SCHEDULE

The age class specific mortality rates acting over some span may be estimated from a feasible mark and recapture program. Consider, for example, an age stratified extension of the multiple mark-recapture method of Seber (1965) and Jolly (1965). Minimally, this requires two marking episodes separated by the span of one age class, followed by one census an equivalent interval later.

At time t a set of individuals are marked and released in the population, and the ages and numbers of individuals are recorded. Let the number of individuals marked and released, classified in age categories spanning a unit time interval, be given by the vector $\underline{v}_t$.

At time t+1 a second set of individuals is marked in a manner that will be distinguishable from the first, and aged and released. Let the number of individuals, by age class, in this release be given by the vector $\underline{v}_{t+1}$.

It is not necessary that the distribution of ages in either of these sets of marked individuals be the same, or that they be representative of the actual age distribution in the population. Neither must all age classes be represented. Let us assume that age classes a through z are reliably sampled by the available gear, and that these age classes are included in the release.

At time t+2 the population is sampled in a fashion that will not bias the relative capture rates among age classes a through z. The numbers of individuals marked at time t+1 and captured at time t+2, grouped according to their current ages at time t+2 are recorded in the vector $\underline{u}_{t+1}$. The numbers of individuals recaptured from among those marked at time t are recorded similarly in vector $\underline{u}_t$.

The fraction of the $v_{x,t+1}$ individuals marked at age x at time t+1 that are recaptured as one age class older individuals at time t+2, and tallied as $u_{x+1,t+1}$, depends on the efficiency of the recapture procedure (as influenced by such factors as effort and dispersal) and on the mortality q_x prevailing between age class x and x+1. Thus

$$u_{x+1,t+1} = V_{x,t+1} \; s \; (1 - q_x) \; , \quad (3)$$

and

$$u_{x+2,t+1} = V_{x+1,t+1} \; s \; (1 - q_{x+1}) \; , \quad (4)$$

where s is an unknown proportionality constant depending on the recapture efficiency.

The fraction of the $v_{x,t}$ individuals marked at age x at time t and recaptured as two age class older individuals at time t+2, and tallied as $u_{x+2,t}$,

depends on the mortality over age classes x and x+1 as
well as the recapture effeciency. Thus

$$U_{x+2,t} = V_{x,t} \, s \, (1 - q_x)(1 - q_{x+1}) \quad . \tag{5}$$

Since both sets of marked individuals are recaptured
in the same sampling, there is no question that the effort
component of the recapture efficiency is the same in
equation (5) as in (3) and (4). We also require, however,
that the dispersion component be the same in order for
s to have the same value for both sets of equations. This
may be assured in a variety of ways.

If mixing of the released individuals into the
population is essentially complete after one time interval
has elapsed, then both releases will be equally dispersed
by the time of the recapture. This would be rather too much
to assume for almost any pelagic stock (and it is just this
problem that poses a major obstacle to the use of mark
and recapture programs for estimates of population size).
In special cases one might be able to estimate the diference
in dispersal in the two releases and correct accordingly,
but this information will not usually be available.

Alternatively, if the recapture program is conducted over a sufficiently extensive area, we may disregard dispersal altogether, for then the individual's position becomes irrelevant to the probability of recapture. This will hold true for the raw recapture data if the recapture effort is well distributed over stations. If certain regions are over-represented in the recapture program, a measure of relative effort per station may readily be used to normalize the recapture data. Note that now we need not assume that the marked individuals are perfectly mixed into the population, for we are concerned only with the relative recapture rates of the different age classes, and not with the capture rates of marked relative to unmarked individuals.

Dividing equation (5) by (4), we obtain after rearrangement

$$(1 - q_x) = \frac{U_{x+2,t} \, V_{x+1,t+1}}{U_{x+2,t+1} \, V_{x,t}} , \quad (6)$$

where all the terms on the right hand side of the equation are measured quantities. Thus, the mark and recapture program provides sufficient information to calculate

the mortality rates from one age class to the next, for
all age classes from a through (z-1).

Now, for all integer valued age classes older than
a, the survivorship is given by

$$\ell_{(x)} = \ell_{(a)} \prod_{j=a}^{x-1} (1 - q_j) \qquad (7)$$

Thus for any integer valued age between a and z, we can, on
the basis of the mortality data obtained from the mark
and recapture program, calculate the relative survivorship.
The value of l(a) remains an unknown proportionality
constant.

CALCULATION OF THE GROWTH RATE

The relative abundances of age classes a through z
in the sampling at time t+2 may be used as an estimate of
the relative age distribution over these age classes. Let
$\underline{n}$ be the vector of abundances of individuals, by age class,
in the sampling. Then, for integer valued ages between a
and z, if the sampled population is in stable age distribution
we may write

$$C^{*}(x) = g \, n_x \quad , \tag{8}$$

where g is an unknown proportionality constant which
absorbs the ratio between the segment of the population
that is between ages a to z and the total population,
and includes also a normalization factor.

Substituting equations (8) and (7) into equation (2)

$$g \, n_x = \frac{e^{-rx} \, \ell_{(a)} \prod_{j=a}^{x-1} (1-q_j)}{\int_0^\infty e^{-ry} \, \ell_{(y)} \, dy} \tag{9}$$

Rearranging, and taking the logarithm of both sides yields

$$\ln\left(\frac{n_x}{\prod_{j=a}^{x-1}(1-q_j)}\right) = -rx - \ln\left(\frac{g}{l(a)}\int_0^\omega e^{-ry}l_{(y)}dy\right).$$

$$(10)$$

Values of the left hand side of the equation, for
integer valued ages between a and (z-1), may be calculated
from the mark and recapture results. Values in the
logarithm on the right hand side of the equation are
independent of x. Accordingly, regression of values for
the left hand side of the equation against age will yield
a straight line with a slope of negative r, the parameter
we wished to calculate; the various unknown scale factors
get collected into one product, the value of which is
determined by the intercept.

Departures of the actual age distribution from the
stable age distribution will be detectable as departures
of the actual points from a straight line. Errors in the
calculation of mortalities arising from year to year
variations in mortality rates of all age classes will
compound geometrically in the denominator of the left
hand side of the equation and the error will thus appear
in the slope of the regression, but will not, in itself

distort the shape of the line. If, for example, all the apparent survival fractions, $(1-q_j)$ are off by a multiplicative factor ϵ, the calculated r will be off by an additive factor $\ln(\epsilon)$.

Of course, actual fluctuations in mortality rates over time will cause the population age distribution to deviate from the stable age distribution. If these fluctuations are short term, and are random in time, they will merely introduce a scatter about the straight line without throwing off the slope by much. Systematic trends in time, however, will cause the shape of the line to change. Keyfitz, Nagnur and Sharma (1967) discuss the analysis of age distributions in populations where the per capita birth rate has been changing either linearly or quadratically over time.

CONSTANT MORTALITIES

If, for the span of age classes considered, the
mortality rate does not change with age, a further simpli-
fication is possible. Then, all the q_j are identical, so
equation (10) becomes

$$\ln(n_x) = x\left(\ln(1-q)-r\right) - \ln\left(\frac{a \ln(1-q)}{l(a)}\int_0^\infty e^{-ry} l_{(y)}dy\right)$$

$$= x \ln\left(\frac{1-q}{\lambda}\right) - \ln\left(\frac{a \ln(1-q)}{l(a)}\int_0^\infty e^{-ry} l_{(y)}dy\right)$$

$$(11)$$

where q is the age independent mortality rate applying to
this span of age classes, and λ is the annual factor of
increase, obtained as e^r. Equation (11) shows that the
regression of $\ln(n_x)$ against x should yield a straight
line, with a slope of $\left(\frac{1-q}{\lambda}\right)$.

Here the linearity displayed in the actual data will
serve as a test both of the assumption of stable age distribution
and of the assumption of constant mortality rates within

the span of ages considered. Such plots have been utilized
in wildlife management as a means to estimate the mortality
rate q where the factor of increase is assumed equal to
one, corresponding to zero population growth (Chapman and
Robson, 1960). Obviously, if q were estimated independently,
we need not make assumptions about λ , but instead would
calculate it from the observed slope.

When neither λ nor q are known, there is still useful
information residing in the measurable ratio $\left(\frac{1-q}{\lambda} \right)$ obtained
from the log plot. Let us assume that the span of ages for
which the constant mortality rate applies includes the ages
within which all (or almost all) reproduction occurs. Assume
further that fecundity during these ages is fairly constant
with age, so that we may utilize the geometric series
approximation to the characteristic equation of the Leslie
matrix:

$$\lambda^{\alpha} - (1-q)\lambda^{\alpha-1} - m\ell_{\alpha} = 0$$

$$(12)$$

where m is the mean fecundity, α is the age at first
reproduction, and ℓ_{α} is the survivorship until the age at
first reproduction. This may be rearranged to

$$\lambda = \left[\frac{M \ell_\alpha}{1 - \left(\frac{1-q}{\lambda} \right)} \right]^{\frac{1}{\alpha}}$$

(13)

Thus we see that the ratio $\left(\frac{1-q}{\lambda} \right)$ is utilizeable even if
we cannot initially separate the mortality from the growth
component. Data on α and m are often available, but
admittedly it will be a rare situation where ℓ_α is
known with sufficient precision to complete the calculation
is equation (13). Nevertheless, even the ability to
compute values of the factor of population increase
contingent upon an unknown value of survivorship to first
reproduction may be useful: we can generally at least
set plausible limits on the value for ℓ_α and this will then
yield the plausible range of λ from equation (13). For
example, in assessing the sustainability of a harvest, this
method could be used to determine whether the presumed
replacement yield is within the limits of known constraints
on the dynamics of the population.

II. ESTIMATING LIFE HISTORY PARAMETERS FROM TRANSIENT AGE DISTRIBUTIONS

When a populations total birth rate is not following an exponential trajectory through time, whether because of variations in the realized vital rates, or because of some past perturbations of the age structure, the simple formula, eq (2), relating survivorships to elements of the age distribution vector, via the population growth rate, no longer applies. It will remain true, however, that the age distribution provides information concerning the vital rates; but that information is now more difficult to extract.

Two important theorems bear on the interpretation of age distributions. The classic proof of convergence of age distributions under a constant life table (Sharpe and Lotka, 1911) implies that once this convergence has been achieved, the age distribution can be calculated from just the fecundity and survivorship schedules. This relationship gives rise, for example, to our equation (2). The second theorem proves that populations experiencing the same history of time-varying life tables will converge in their age distributions, though these age distributions will themselves be time dependent (Lopez, 1961).

The second theorem is less widely known than the
first among ecologists and wildlife managers, though it
has an important consequence that is identical to a
feature that follows from the classic theorem. In fact,
the classic theorem is a special case of Lopez' proof.
The noteworthy property demonstrated in either case is
that the initial age distribution becomes essentially
irrelevant to the current age distribution, as time
progresses. Thus, even with time varying life tables,
a sufficiently long time sequence of vital rates--age
specific fecundity and mortality schedules--will
suffice to determine the age distribution.

Accordingly, mathematical manipulation of the age
distribution must reveal some constraints on the sequence
of vital rates. If enough properties of the sequence of
vital rates are already known, the constraints provided
by the analysis of the age distribution can finally yield
a solution for the vital rates themselves.

The convergence properties exhibitted by continuous
time population models do not invariably carry over to the
analogous discrete time Leslie matrix models. The precise
condition for convergence of age structure in the
matrix version is that there be no common divisor greater
than one for the set of ages at which reproduction

occurs (Lopez, 1961; Sykes, 1969). This condition is almost certain to be met in annual Leslie matrix models for long lived repeatedly breeding organisms. For example, if any two successive age classes are reproductively active, these two ages must be relatively prime.

CONVERGENCE OF AGE DISTRIBUTION UNDER A VARIABLE HARVEST

Consider a population subject to schedules of natural mortality and fecundity which are encapsulated in the Leslie matrix, A, which is constant over time. In the usual notation

$$A = \begin{pmatrix} m_1 & m_2 & m_3 & \cdots & m_{\omega-1} & m_\omega \\ p_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & p_3 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & p_{\omega-1} & 0 \end{pmatrix} \qquad (14)$$

where m_i is the effective fecundity of age class i, and p_i is the probability of survival from age i to age i+1.

Beginning at some time 0 imagine that an annual harvest is instituted, and that the number of individuals harvested in year t, by age class, is recorded in the vector $\underline{h}_t$, which may vary from year to year. Thus the realized vital rates in year t, which project the population vector to time t+1, are encapsulated in the matrix $\tilde{A}_t$, where the fecundity vector is related to the fecundity vector of A, the harvest vector $\underline{h}_t$, and the

population vector $\underline{n}_t$, according to

$$\overset{u}{n}_{i,t} = u_i \left(1 - \frac{h_{i,t}}{\sum\limits_{j=1}^{w} u_j n_{j,t}} \right)$$

$\qquad\qquad (15)$

and the survival rate vector is related to the survival rate vector of A, the harvest vector, and the population vector, according to

$$\overset{u}{p}_{i,t} = \frac{p_i n_{i,t} - h_{i+1,t}}{n_{i,t}}$$

$\qquad\qquad (16)$

assuming that the harvest is taken at the end of each time interval.

We might project the population vector from time 0 to time t by multiplying by the succession of matrices:

$$\underline{n}_t = \overset{u}{A}_{t-1} \overset{u}{A}_{t-2} \overset{u}{A}_{t-3} \cdots \overset{u}{A}_0 \underline{n}_0$$

$\qquad\qquad (17)$

but this would be very awkward, owing to the involvement of all the intermediate population vectors in the succession of matrices, and it seems unlikely that the initial population vector will be known in many important practical cases.

Alternatively, we may note that in each time interval the population grows according to A and then is reduced by the current harvest:

$$\underline{n}_{k+1} = A\underline{n}_k - \underline{h}_k \tag{18}$$

Thus, applying this formula recursively from time 0 to t, we have

$$\underline{n}_t = A^t \underline{n}_0 - \sum_{k=0}^{t-1} A^{(t-k-1)} \underline{h}_k \tag{19}$$

Intuitively, we see that the final population vector is a sum of as many component vectors as there were harvests. The initial vector is projected through t applications of the constant matrix of natural vital rates, A. We subtract from this vector the respective harvests, where the final effect of each harvest is given by projecting this phantom population through the number of time units that elapsed since it was taken.

We observe that the initial population vector figures only once in the expansion in equation (19), and that is in the term where the constant projection matrix is applied to

it t times. From classical stable population theory, we know that as t becomes sufficiently large, the vector projected by $A^t \underline{n}_o$-will converge on the stable age distribution associated with the matrix A, regardless of the age distribution in the initial vector. Formally

$$A^t \underline{n}_o = v \underline{c}^*$$

(20)

where $\underline{c}^*$ is the stable age distribution of A (it is the eigenvector associated with the dominant eigenvalue of A) calculated as the discrete time equivalent of equation (2) , and v is a scalar multiplier. Only the value of the scalar is determined by the initial population: it may be calculated as the summed reproductive value in the initial population, multiplied by λ^t , where λ is the dominant eigenvalue of matrix A.

Thus the w variables that would be unknown if the initial population were unknown, collapse to a single unknown in their effect on the present population vector:

$$\underline{n}_t = v \underline{c}^* - \sum_{k=0}^{t-1} A^{(t-k-1)} \underline{h}_k$$

(21)

CALIBRATING THE LIFE TABLE FROM THE AGE DISTRIBUTION OF
A HARVESTED POPULATION

It is often very difficult to obtain estimates of
population size, for the reasons discussed in section I,
but reliable estimates of the relative abundances of
several age classes may be readily available. Similarly,
programs of mark and recapture may reveal patterns of
relative survival rates, and examination of the frequencies
of pregnancy in aged samples may reveal relative fecundities
(but not the crucial rate of survival from birth until
the first census).

Let us then imagine that we have trustworthy estimates
for the relative magnitudes of all the elements of a Leslie
matrix of natural vital rates, so the entire Leslie matrix is
known except for a proportionality constant β. Thus where
B is the matrix of relative rates, the actual Leslie matrix
A is given by

$$A = \beta B$$

(22)

And let the population vector $\underline{n}_t$ be known only to a
proportionality constant g; and then perhaps only for
some age classes:

$$\underline{n}_t = g\,\underline{c}_t$$

$$(23)$$

where the vector of age distribution is available from data, but the multiplier g is unknown.

For a number of commercially cropped populations we have good estimates of the history of the magnitudes and age compositions of the harvests. Indeed, the estimates of harvest may be the most secure knowledge we have about many of these populations.

Substituting in equation (21) we have

$$g\,\underline{c}_t = v\,\beta^t\,B^t\,\underline{c}^* - \sum_{k=0}^{t-1} \beta^{(t-k-1)}\,B^{(t-k-1)}\,\underline{h}_k$$

$$(24)$$

Now, the stable age distribution $\underline{c}^*$ must be the eigenvector of A, but since multiplying a matrix by a scalar does not alter its eigenvectors, $\underline{c}^*$ may be calculated simply as the dominant eigenvector of the known matrix B.

Thus, equation (24) contains but three unknowns: these are the scalars g, v, and β. If the estimated values for the age distribution at time t are reliable for at least three age classes, then we will have enough equations to solve for the unknown, provided the equations are

linearly independent. If there has been at least some

harvest, this condition is satisfied.

The form of the solution is easily seen, for example,

by treating $v\beta^t$ as one unknown and g as another. Simple

algebraic manipulation of three equations of the form

$$g c_{i,t} = (v\beta^t) a_i - (b_{0,t}\beta^{t-1} + b_{1,t}\beta^{t-2} + \cdots + b_{t-1,t}) \tag{25}$$

(where a_i , $c_{i,t}$ and all the b's are knowns) will leave us

with one polynomial equation of order t-1 in β , and with

the other unknowns canceled out.

From Descartes' rule of signs, there will be only

one real positive root for the polynomial. This root may

readily be obtained by a simple iterative technique,

such as Newton's method, yielding the value β that completes

our knowledge of the life table. Where more equations

than necessary are in hand, the overdetermined system may

be "fit" statistically, much as we fit a straight line

to the data in the method of section I, rather than

simply solving for the unknown. The degree of fit will

then provide some estimate of internal consistency.

In general, the missing parameter in an otherwise
known Leslie matrix probably will not be a multiplier
that is applied to the entire matrix. This case was
explored here merely for the algebraic convenience which
permitted explicit formulation of equation (24). In real
cases it is more likely that survival rates will be
known, but that a proportionality constant to be applied
to the fecundity vector, or perhaps to the survival rates
of just the immatures, will be missing. Then the mathematics
will be less esthetic, and the steps past equation (21)
will require recourse to iterative computer techniques,
but the logic of the solution remains the same.

The key feature is the step in equation (20) which
replaces the w unknowns of the initial population vector
with one, freeing up the remaining equations to determine
unknowns which are of greater interest. Indeed, if the
relative abundances of a sufficient number of age classes
in c_t can reliably be estimated, it becomes possible,
by this means, to solve simultaneously for more than one
missing parameter of the Leslie matrix.

Two assumptions of the calculations outlined in
this section may not fully be met in some applications.
One is that the time span of the history of the harvests

is long enough to guarantee the convergence of the
contribution of the initial population vector to stable
form. In practise, it appears that convergence is quite
rapid, most especially since the initial vector may not
be terribly far from stable form. Should any doubts
arise on this score in a particular application, the
validity of the identity in equation (20) may be explored
in simulations by initializing $\underline{n}_o$ with a plausible
range of age distributions, and examining the divergence
of $A^t\underline{n}_o$ from some multiple of $\underline{c}^*$. The propagation of
the resultant error through the ensuing calculations is
easily traced, allowing some estimates of confidence
intervals on the final parameter that is solved for.

A second problem arises if the natural component
of the vital rates, that is A, varies significantly over
time. Provided enough is known about the time sequence
of the $\left\{A_t\right\}$, the essential point, that the initial
age distribution need not be known remains true. That
is, the formulation, while not so simple as equation (19),
still will include the initial population vector in
a single term only, namely $A_{t-1}A_{t-2}\cdots\cdots A_1A_o\underline{n}_o$, and
the age distribution of the ensuing contribution is
essentially a property of the sequence of A's rather than
of $\underline{n}_o$, as per Lopez' proof. Accordingly, carefully

constructed computer algorithms, even in this less favorable situation, can extract information about vital rates from the current age distribution and a record of harvests.

If there is suspicion of a slow systematic change in A_t , as would ccur with density dependent changes if the populations size is gradually changing, it may be advantageous to truncate the population history so as to minimize the degree of change or to restrict the change to a period when the consequences on A_t are best understood. Then the new time 0 will not be the onset of exploitation but simply the starting time of the analysis sequence. Since the initial population is in any case treated as an unknown, this shift is unimportant. What is important is that the sequence remain long enough for convergence of the initial population vector's contribution to the age distribution. Reflection on this strategy shows, incidentally, that the most recent harvest records will tend to be most critical in the calculations.

RELATED METHODOLOGIES

The estimation of population parameters from a time sequence of harvests may be carried out via techniques of comparing observed and predicted age compositions in the harvests themselves (Doubleday, 1975; Free and Beddington, 1979). In these analyses, no a priori knowledge of actual age distributions in the population is presumed. Instead, where there are k recorded harvests of a w age class population, the harvests constitute a set of $k \cdot w$ observations, which are fit by adjusting a set of population parameters and a set of selectivity parameters. The selectivity parameters for a given year form the vector that converts the age distribution of the harvest to the age distribution in the actual population. If the selectivity changes over time, in response for example to market and gear changes, there can, at worst, be a total of $k \cdot w$ selectivity parameters, which would leave no degrees of freedom for the fitting of population parameters. In practise, it is assumed that selectivity is constant for some period of time, or for some blocks of age classes, thus restoring a more favorable relation between the number of observations and the number unknowns.

If these assumptions about the selectivity of the historical harvests are valid, this technique of fitting the harvest record is superior to the method developed in this paper in that each harvest contributes additional

degrees of freedom, whereas we propose only one final

comparison of an observed versus a predicted age distribution.

If, on the other hand, the assumptions about historical

harvest selectivities seem dubious (or inaccessible to

evaluation), the method proposed in this paper permits the

information content of the historical harvests to contribute

to the calculation, without incorporation of objectionable

assumptions, provided we can obtain an independent estimate

of the final age distribution (equivalent to obtaining

an independent estimate of the harvest selectivities applying

to the final harvest).

It has been observed in numerical exploration of the

method of simulataneously fitting selectivities and population

parameters that the residual is insenesitive to the age

composition of the initial population (Free and Beddington,

1979). We can now understand this as a consequence of the

same convergence propoerty noted in equation (18). The

similarity of the underlying mathematics in the two techniques

suggests a hybrid approach for situations where there is

a long historical record of harvests alone, and a shorter

more recent record of information (such as independent

assessments of actual population composition) bearing on

selectivities. Then we could use the theoretical convergence

argument developed in this paper to get the population from

the begining of the period of recorded harvests to the onset

of the period of availability of selectivity or age composition data, and begin accumulating residuals at that point after the manner of models that attempt to fit selectivity parameters as well.

CONCLUSIONS

Age distributions contain information about the history of birth and death rates in the population. Where the per capita age specific rates are independent of history, the age distribution converges to a characteristic form which may be calculated directly from the life table. This dependence, alternatively, permits computation of some life table parameters from the observed age distribution. Two examples are presented to illustrate the calculation of population growth rate from some minimal data including a partial age distribution.

Where the vital rates are history depedent, the age distribution of the population depends on that history , though certain convergence properties remain. For the case where the history dependence arises via a known history of time varying harvest (known only as absolute rather than per capita rates) we develop a model which utilizes the consequent partial convergence to calculate life table parameters from the record of harvests and a final observed partial age distribution. This method is related to another technique of sequential population analysis which proceeds via comparison of observed and expected harvest structures. The latter technique is preferable where harvest selectivities are understood; our new method is the only option where the history of harvest selectivities is questionable, but one

independent assessment of the population age distribution
is available. In the interdediate situation where the record
of harvest information is longer than the record of selectivity
information, a hybrid analysis can be constructed.

REFERENCES

Bourgeois-Pichat, J. 1957. Utilization de la notion de
    population stable pour mesurer la mortalité et
    la fécondité des populations des pays sous-développes.
    Bull. Inst. Internat. Stat. (Actes de la 30e Session).

Chapman, D.G., and G.I. Murphy. 1965. Estimates of
    mortality and population from survey-removal
    records. Biometrics 21:920-935.

Chapman, D.G., and D.S. Robson. 1960. The analysis of a
    catch curve. Biometrics 16: 354-368.

Coale, A.J. 1972. The growth and structure of human
    populations. Princeton Univ. Press. Princeton, N.J.

Doubleday, W.G. 1975. The least squares approach to
    analysing catch at age data. Int. Comm. Northwest
    Atl. Fish. Res. Doc. 75/35.

Free, C.A., and J.R. Beddington. 1979. A population
    assessment of the Northwest Atlantic harp seal stock
    based on an estimation technique comparing observed
    and expected age distributions of catch. IUCN & WWF
    International Workshop on Biology and Mnagement of
    Northwest Atlantic Harp Seals, WP1, pp517-550.

Gulland, J.A. 1965. Estimation of mortality rates. Annex to
Arctic Fisheries Working Group Report. ICES, C.M. 1965,
Doc. 3.

Gulland, J.A. 1969. Manual of methods for fish stock
assessment. Part 1. Fish population analysis.
FAO Man. Fish. Sci.

Jolly, G.M. 1965. Explicit estimates from capture-recapture
data with both death and immigration--stochastic
model. Biometrika 52:225-247.

Keyfitz, N. 1968. Introduction to the mathematics of
population. Addison-Wesley. Reading, Mass.

Keyfitz, N., D. Nagnur and D. Sharma. 1967. On the
interpretation of age distributions. J. Am.
Stat. Assoc. 62:862-874.

Lopez, A. 1961. Problems in stable population theory.
Office of Population Research. Princeton, N.J.

Lotka, A.J. 1939. A contribution to the theory of self-
renewing aggregates, with special reference to
industrial replacement. Ann. Math. Stat. 10:144-161.

Pope, J.G. 1972. An investigation of the accuracy of
    virtual population analysis using cohort analysis.
    Int. Comm. Northwest Atl. Fish. Res. Bull. 9:65-74.

Seber, G.A.F. 1965. A note on the multiple-recapture census.
    Biometrika 52: 249-259.

Seber, G.A.F. 1973. The estimation of animal abundance.
    Griffin. London.

Sharpe, F.R., and A.J. Lotka. 1911. A problem in age
    distribution. Philos. Mag. (Ser. 6) 21:435-438.

Sykes, Z.M. 1969. On discrete stable population theory.
    Biometrics 25:285-293.