THE MATHEMATICS
OF
STABLE POPULATION THEORY
AND
ITS GENERALIZATION

PREPARED

BY

JOSEPH ANTONY MAKOTEKU OTTIENO

OF

MATHEMATICS DEPARTMENT

AND

POPULATION STUDIES

&

RESEARCH INSTITUTE
UNIVERSITY OF NAIROBI
P.O. BOX 30197
NAIROBI, KENYA

This project was initiated when the author was a Post-Doctoral Fellow at the Office of Population Research, Princeton University, New Jersey in the year 1981/82.

The fellowship was granted by the generosity of

THE FORD FOUNDATION

through

THE INSTITUTE OF INTERNATIONAL EDUCATION

CONTENTS

CHAPTER I

MEASUREMENTS OF MORTALITY AND FERTILITY

1.1 RATIO, PROPORTION AND RATE

1.1.1 Ratio

A ratio results from dividing one quantity by another. It summarises the arithmetical relationship between two characteristics which can be counted in a population. When the number of occurrences are 'a' and 'b', then the ratio r is defined as

$$r = \frac{a}{b} \quad \text{or} \quad r = \frac{b}{a} \qquad (1.1)$$

Another way of expressing

$$r = \frac{a}{b}$$

is simply a : b

Thus the calculation of a ratio is intended to provide further meaningful information concerning the behaviour of two characteristics. By calculating a ratio, the numerator is in some ways adjusted with respect to the denominator which becomes a reference value and leads to an expected value for the numerator.

To eliminate decimal points, ratios are often multiplied by a power of 10. Thus the ratio r becomes

$$r = \frac{a}{b} \times 10^k \qquad (1.2)$$

where k is a non-negative number.

For

$$k = 0,$$

the expression is a simple ratio with unity as the basis of comparison.

For

$$k = 2$$

the expression is a percentage.
If $k = 3$ it is per thousand and for $k = 6$ the basis is a million.

The following are examples of a ratio:

Example 1:

$$\text{Sex ratio} = \frac{\text{number of males}}{\text{number of females}}$$

in a given population

Example 2:

$$\text{Fetal death ratio} = \frac{\text{number of fetal deaths}}{\text{number of live births}}$$

in a given year for a given a population. This expresses the number of fetal deaths as compared to live births in the same population.

Remark: The two examples given above, show how a ratio is used in comparing the frequencies of two mutually exclusive classes.

Example 3: There are 21000 operations and 30 doctors in a hospital in a given year.

The crude numbers alone tell us something about the work load of the hospital and about the staffing.

But an administrator who wishes to consider the expansion of the hospital and how many additional doctors he might need would have to do further analysis. He could start by calculating the ratio

$$r = \frac{a}{b}$$

$$= \frac{21000}{30}$$

$$= 700$$

i.e., 700 operations per doctor per year.

This would suggest that one doctor was required for 700 operations. In this case, k = 0, and the ratio relates to the work of one doctor.

Conversely, he might be interested in comparing different hospitals with different numbers of operations, then in order to adjust for these different numbers he would calculate the ratio

$$r = \frac{b}{a}$$

$$= \frac{30}{21000}$$

$$= 0.00143$$

i.e., 0.00143 doctors per operation.

This figure seems meaningless, so we would put k = 4 so that we hav

$$r = \frac{b}{a} \times 10^4$$

$$= 14.3$$

i.e. 14.3 doctors per 10,000 operations or 143 per 10,000. This value of r is now a useful tool for comparing different hospitals with each other, adjusting for the differences in work load.

Example 4: A ratio is not limited to the relationship between two counts but extends to all measurable variables. For example, if the mean height of a sample of women is 165 cm and the mean height of a sample of males is 175 cm, then the male to female ratio of height is

$$r = \frac{175}{165} = 1.06$$

which implies that on the average, men are 6 per cent taller than women.

1.1.2 Proportion

Proportion is a special type of ratio in which the numerator is a part of the denominator, i.e.,

$$p = \frac{a}{a+b} \qquad (1.3)$$

If the numerator and denominator are integers and represent frequencies of certain events, then p is a relative frequency. For example

$$\frac{\text{number of males}}{\text{number of males + number of females}}$$

gives the proportion (relative frequency) of males in a given community.

In a large population, proportion may determine the probability of a certain event; in a sample (experiment) proportion can be used as an estimate of probability of an event. For example, the quantity

$$\frac{\text{number of fetal deaths}}{\text{number of fetal deaths + number of live births}}$$

in a given population for a given year is clearly a proportion. It estimates the probability that a fetus might die before it is born.

Generally, the numerator and the denominator in a/(a+b) do not need to be integers. They can be measurable quantities such as weights, lengths, volume etc. In such cases proportions are also often called <u>fractions</u>. For example, in a chemical analysis, the mass of a given component can be expressed as a fraction of the total weight of the compound. Percentage is a proportion or a fraction per hundred units. This proportion p corresponds to 100p percent.

1.1.3 Rates

Ratios and proportions are useful static summary measures of phenomena that occurred under certain conditions. The concept of rate is associated with the dynamics of phenomena such as chemical reactions, growth, birth, death, spread of epidemics etc. Generally, rate can be defined as a measure of change in one quantity (y) per unit of another quantity (x) on which y depends. Usually the independent variable (x) is time, although it might represent some other physical quantities such as temperature or pressure. For convenience we mostly confine ourselves to processes depending on time, and denote time by t rather than by x.

For definitions of absolute rate, relative rate, central rate and rates for repetitive events, the reader is referred to Elandt Johnson and Johnson. We shall confine ourselves with the rate of incidence which is defined as the number of events that occur within a given time interval over the number of members of the population who exposed to the risk of the event during that same time interval.

Specifying the number of persons exposed to risk in the denominator is an important refinement. If we were studying mortality over one-year period, we should note that a person who died before the year ended was not exposed to risk for the whole year and neither was a child who was born halfway through the year. People who moved to a country only one month before the year ended, were not exposed to the risk of dying for the whole year either. The concept of 'person-years lived' is often used to specify the population exposed to the risk of an event.

Let N_T denote the number of individuals ever observed during a period of T years. Let T_j denote the length of the period (in years) during which an individual (j) was under observation, that is, exposed to risk of being observed to die. Then the sum of lengths of such periods of exposure

$$A_T = \sum_{j=1}^{N_T} 1.T_j \qquad (1.4)$$

gives the total amount of person years (analogous to mass-time) exposed to risk.

As an example let T = 3 years. Suppose that $N_T = 10$ individuals were observed during this period further following lengths of time T_j : 2.3, 1.5, 2.8, 2.5, 3.0, 1.8, 2.7, 2.5, 3.0, 3.0 years.

Thus the number of person-years is given by

$$A_T = 2.3 + 1.5 + 2.8 + 2.5 + 3.0 + 1.8 + 2.7 + 2.5 + 3.0 + 3.0$$
$$= 24.3$$

1.2 Measures of Mortality

Crude Death Rate (CDR)

$$CDR = \frac{\text{Deaths during a specified period}}{\text{Person-years lived during the period of population at risk}}$$

$$= \frac{D}{P} \qquad (1.5a)$$

Conventionally rates are expressed per 1000 people. So we have

$$CDR = \frac{D}{P} \times 1000 \qquad (1.5b)$$

P is approximated by the mid-year population.

Specific Death Rates

A population may be divided into sub-populations according to one or more factors of classification such as age, sex, marital status, occupation, urban/rural incidence, duration of marriage etc. These are called specific death rates. In other words, a specific death rate is simply one that refers only to some sub-group in the population. If the sub-populations are individuals in different age groups the resulting rates are called age-specific death rates.

Age Specific Death Rate $({}_nM_x)$

$$ {}_nM_x = \frac{{}_nD_x}{{}_nP_x} \tag{1.6} $$

$$ = \frac{\text{Deaths of persons aged between } (x,\ x+n)}{\text{Person-years lived by the population in that group.}} $$

The relationship between Crude Death Rate and Age-Specific Death Rate can be shown as follows:-

$$ CDR = \frac{D}{P} $$

$$ = \frac{\sum_x {}_nD_x}{\sum_x {}_nP_x} \tag{1.7} $$

But

$$ {}_nD_x = {}_nM_x \cdot {}_nP_x $$

from (1.6).

Therefore

$$ CDR = \frac{\sum_x {}_nM_x \cdot {}_nP_x}{\sum_x {}_nP_x} \tag{1.8} $$

For the continuous case

$$ CDR = \frac{\int_0^w \mu(x)\, P(x) dx}{\int_0^w P(x)\, dx} \tag{1.9} $$

where $\mu(x)$ is the death rate at exently age x i.e.,

$$ \mu(x) = \lim_{n \to 0} {}_nM_x. \tag{1.10} $$

The greatest age reached is denoted by w.

Let

$$ {}_nW_x = \frac{{}_nP_x}{\sum_x {}_nP_x} \qquad (1.11) $$

Then

$$ CDR = \sum_x {}_nM_x \frac{{}_nP_x}{\sum_x {}_nP_x} $$

$$ = \sum_x {}_nM_x \,.\, {}_nW_x \qquad (1.12) $$

${}_nW_x$ is a measure of the relative size of the sub-group under consideration such that

$$ \sum_x {}_nW_x = 1 \qquad (1.13) $$

In other words the crude death rate is merely a weighted average of the specific death rates. Thus a change in the crude death rate may result from a change in the relative sizes of the component sub-populations subject to the restriction $\sum_x {}_nW_x = 1$ without any change in the specific death rates measuring the intensity of death per time unit in the sub-populations.

So crude death rate depends on age composition. A young population has low CDR while an old population has high CDR.

Infant Mortality Rate

For demographic statistical purposes, all children under one year of age are considered infants and so the term "infant mortality" refers to mortality among children of less than one year of age.

The infant mortality rate may be defined as the number of infant deaths that occur per thousand live births in any population in one calendar year. Thus if D_0 is the number of deaths occuring for those under one year and B is the number of live births in the same year in the same community, then the infant mortality rate is defined as:

$$IMR = \frac{D_o}{B} \times 1000 \tag{1.14}$$

Remark: IMR is not truly a rate nor a proportion. It is a ratio for which the time over which it is recorded need not be specified so long as births and infants deaths are recorded simultaneously.

1.3 Measures of Fertility

Fertility is a measure of production of live births. Fecundability is the capacity of bearing a live birth.

Crude Birth Rate (CBR)

$$CBR = \frac{\text{Births to a population in a period}}{\text{Person - Years lived}}$$

It measures the fertility of the population as a whole, rather than of that segment of the population biologically capable of bearing children.

Due to limitations or distortions of the CBR, we need to use other measures of fertility.

General Fertility Rate (GFR)

$$GFR = \frac{\text{Number of births in a period}}{\text{Total number of women of child-bearing age-span}}$$

Age Specific Fertility Rate (ASFR)

$$ASFR = \frac{\text{Births in a period to women in a specific age interval}}{\text{Person-years lived in this age interval}}$$

$$= \text{births per woman per year in the specific age interval}$$

Total Fertility Rate (TFR)

TFR = Average number of births per woman that would occur to a hypothetical cohort of women subject through its life to the given fertility schedule.

= Sum of ASFR over all the age intervals.

Notations

Let

f_i = Births per woman-year in the ith age interval

w_i = Number of women in the ith age interval

N = Total population of both sexes

W = Total number of women in the child bearing age.

B = Total number of births

$$C_i = \frac{w_i}{N}$$

and

$$C_i^* = \frac{w_i}{W}$$

Therefore

$$CBR = \frac{B}{N}$$

$$= \sum_i \frac{w_i f_i}{N} \qquad (1.15)$$

$$= \sum_i c_i f_i \qquad (1.16)$$

$$GFR = \frac{B}{W}$$

$$= \sum_i \frac{w_i f_i}{W} = \sum_i c_i^* f_i \qquad (1.17)$$

$$TFR = \sum_{i=x}^{\beta} f_i \qquad (1.18)$$

where α and β are the lower and upper limits of the child-bearing age span.

Suppose $\alpha = 15$ and $\beta = 50$, then in a 5-year age interval, we have

$$TFR = 5 \sum_{i=1}^{7} f_i \quad (1.19)$$

In the continuous case

$$CBR = \int_{\alpha}^{\beta} C(x) f(x) dx \quad (1.20)$$

$$GFR = \int_{\alpha}^{\beta} C^*(x) f(x) dx \quad (1.21)$$

and

$$TFR = \int_{\alpha}^{\beta} f(x) dx. \quad (1.22)$$

Gross Reproduction Rate (GRR)

The gross reproduction rate is identical to total fertility except that it sums female births only, so that it indicates the total number of daughters that would be born. So GRR is the number of female children per woman subject to given fertility schedule from α to β.

Thus

$$GRR = \Sigma \text{ ASFR (females)}$$

$$= TFR \cdot \frac{\text{female births}}{\text{female births + male births}} \quad (1.23)$$

If for every 100 females there are 106 males, then

$$GRR = TFR \cdot \frac{100}{100 + 106}$$

$$= \frac{TFR}{2.06} \quad (1.24)$$

GRR = 1, implies that the females will be replacing themselves.

<u>Relationship between total fertility rate and general fertility rate.</u>

Let 'a' be a uniform random variable within the interval α and β. So the distribution of 'a' can be written as

$$h(a) = \frac{1}{\beta-\alpha} \quad , \quad \alpha < a < \beta$$

$$= o, \text{ otherwise.} \qquad (1.25)$$

If f(a) and c*(a) are functions of age, then they are also random variables with expectations

$$E[f(a)] = \int_{\alpha}^{\beta} f(a)h(a)da$$

$$= \frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} f(a)da \qquad (1.26)$$

and

$$E[c^*(a)] = \int_{\alpha}^{\beta} c^*(a)h(a)da$$

$$= \frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} c^*(a)da \qquad (1.27)$$

Further

$$E[f(a)\ c^*(a)] = \int_{\alpha}^{\beta} f(a)c^*(a)h(a)da$$

$$= \frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} f(a)c^*(a)da \qquad (1.28)$$

Therefore correlation coefficient between f(a) and c*(a) is given by

$$r_{f.c^*} = \frac{E[f(a)c^*(a)] - E[f(a)]\ E[c^*(a)]}{\sigma_f\ \sigma_{c^*}}$$

i.e.,

$$r_{f.c^*} = \frac{\frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} f(a)c^*(a)da - \frac{1}{(\beta-\alpha)^2} \int_{\alpha}^{\beta} f(a)da \int_{\alpha}^{\beta} c^*(a)da}{\sigma_f\ \sigma_c}$$

$$= \frac{\int_\alpha^\beta f(a)c^*(a)da - \frac{1}{(\beta-\alpha)} \int_\alpha^\beta f(a)da \int_\alpha^\beta c^*(a)da}{(\beta - \alpha)\ \sigma_f\ \sigma_{c^*}} \qquad (1.29)$$

This implies that

$$r_{f.c^*}\ (\beta-\alpha)\sigma_f\sigma_{c^*} = \int_\alpha^\beta f(a)c^*(a)da - \frac{1}{\beta-\alpha} \int_\alpha^\beta f(a)da \int_\alpha^\beta c^*(a)da$$

$$= GFR - \frac{1}{\beta-\alpha}\ TFR \qquad (1.30)$$

using (1.21) and (1.22) and the fact that

$$\int_\alpha^\beta c^*(a)da = 1 \qquad (1.31)$$

Therefore

$$GFR = \frac{TFR}{\beta-\alpha} + (\beta - \alpha)\ r_{f.c^*}\ \sigma_f\ \sigma_{c^*} \qquad (1.32)$$

Net Reproduction Rate

Let m(a) be the proportion of women at age 'a' who bear a female child. Thus the gross reproductive rate can now be defined by

$$GRR = \int_\alpha^\beta m(a)da \qquad (1.33)$$

Let p(a) be the proportion surviving from birth to age 'a'. Thus the proportion of women who survive to age 'a' and bear a daughter is given by

$$\phi(a) = p(a)\ m(a) \qquad (1.34)$$

Net Reproduction Rate (NRR) is defined as

$$NRR = \int_\alpha^\beta p(a)\ m(a)da \qquad (1.35a)$$

$$= \int_\alpha^\beta \phi(a)da \qquad (1.35b)$$

We may call

$p(a)$, the schedule of survival

$m(a)$, the schedule of maternity

and $\phi(a)$, the schedule of net maternity.

If $p(\bar{a})$ is the probability of surviving from birth to the mean age of child hearing, then

$$NRR = \int_{\alpha}^{\beta} p(\bar{a}) \, . \, m(a)da$$

$$= p(\bar{a}) \int_{\alpha}^{\beta} m(a)da$$

i.e.,

$$NRR = p(\bar{a}) \, . \, GRR \, . \qquad (1.36)$$

If $m(a)$ = constant, say k, then

$$GRR = \int_{\alpha}^{\beta} m(a)da$$

$$= \int_{\alpha}^{\beta} k \, da$$

$$= k \, (\beta - \alpha). \qquad (1.37)$$

Other measures of fertility

Completed family size: This indicates for a woman at the end of her reproductive period, the total number of live births she reports to have had.

Parity: The number of children born to a woman in the different age-groups. Thus parity for a woman over 50 years measures the completed family size.

CHAPTER II

THE LIFE TABLE

2.1 Introduction

A life table is a device for exhibiting the mortality history of an artificial population, called a cohort, as it gradually decreases in size until all its members have died.

A more general definition of a life table is that it is a device exhibiting the history of a hypothetical cohort subject to attrition. The attrition could be death, cause of death, marriage dissolution, in-migrants and out-migrants, contraceptive failure, duration of breast-feeding, replacement of buses due to deterioration etc.

For the case of death, we refer to age specific while in contraceptive and migration, we talk of duration specific. A life table can thus be a function of duration for the case of contraceptive failure and the like, while it is a function of age for the case of death. Life tables that deal with age intervals of one year are frequently referred to as complete life tables, whereas those using longer intervals are called abridged life tables.

2.2 Life table functions and their relationships

2.2.1 Survival function $\ell(x)$

Let $\ell(x)$ or ℓ_x be the number of survivors at age x, out of the initial population size $\ell(o)$. The initial group or cohort is called a radix which is set equal to some arbitrary constant, usually the power of 10 such as 1000, 10,000, 100,000 etc. If, however, the radix is set equal to unity, then $\ell(x)$ becomes the probability of an individual surviving up to age x. Otherwise the proportion of survivors at exact age x is

$$p(x) = \frac{\ell_x}{\ell_o} \quad (2.1)$$

The expected number of deaths in (x, x + 1) is given by

$$d_x = \ell_x - \ell_{x+1} \qquad (2.2)$$

Thus the conditional probability that an individual dies at age x + 1 given that he was alive at age x is

$$q_x = \frac{d_x}{\ell_x} \qquad (2.3a)$$

$$= \frac{\ell_x - \ell_{x+1}}{\ell_x} \qquad (2.3b)$$

The probability that an individual at exact age x will not die in (x, x + 1), i.e., will survive beyond age x + 1 is

$$p_x = 1 - q_x \qquad (2.4a)$$

$$= \frac{\ell_{x+1}}{\ell_x} \qquad (2.4b)$$

Thus we can express ℓ_x as

$$\ell_x = \frac{\ell_x}{\ell_{x-1}} \cdot \frac{\ell_{x-1}}{\ell_{x-2}} \cdot \frac{\ell_{x-2}}{\ell_{x-3}} \cdot\cdot\cdot \frac{\ell_3}{\ell_2} \cdot \frac{\ell_2}{\ell_1} \cdot \frac{\ell_1}{\ell_o} \cdot \ell_o$$

$$= p_{x-1}\, p_{x-2}\, p_{x-3} \cdot\cdot\cdot p_2\, p_1\, p_o\, \ell_o \qquad (2.5)$$

Therefore the proportion surviving at age x is

$$p(x) = \frac{\ell_x}{\ell_o}$$

$$= p_o p_1 p_2 \cdot\cdot\cdot p_{x-3}\, p_{x-2}\, p_{x-1}$$

$$= \prod_{y=o}^{x-1} p_y \qquad (2.6)$$

Alternatively, from (2.4b)

$$p_x\, \ell_x = \ell_{x+1}$$

$$p_{x-1}\, \ell_{x-1} = \ell_x$$

$$p_{x-2}\, \ell_{x-2} = \ell_{x-1}$$

$$\vdots$$

etc

So

$$\ell_x = p_{x-1}\, \ell_{x-1}$$

$$= p_{x-1}\, p_{x-2}\, \ell_{x-2}$$

$$= p_{x-1}\, p_{x-2}\, p_{x-3}\, \ell_{x-3}$$

$$\vdots$$

$$= p_{x-1}\, p_{x-2} \cdots p_2 p_1 p_o \ell_o$$

Thus

$$p(x) = \frac{\ell_x}{\ell_o}$$

$$= \prod_{y=o}^{x-1} p_y$$

The probability that a person of exact age x will survive n years, is

$$_n p_x = \frac{\ell_{x+n}}{\ell_x} \qquad (2.7)$$

i.e.

$$_n p_x = \frac{\ell_{x+1}}{\ell_x} \cdot \frac{\ell_{x+2}}{\ell_{x+1}} \cdot \frac{\ell_{x+3}}{\ell_{x+2}} \cdots \frac{\ell_{x+n}}{\ell_{x+n-1}}$$

$$= p_x p_{x+1}\, p_{x+2} \cdots p_{x+n-2}\, p_{x+n-1} \qquad (2.8)$$

The complement,

$$_n q_x = 1 - {_n p_x}$$

$$= 1 - \frac{\ell_{x+n}}{\ell_x}$$

$$= \frac{\ell_x - \ell_{x+n}}{\ell_x} \qquad (2.9)$$

Now consider

n/q_x = Probability that a person aged x may die in the nth year.

= Probability that a person aged x survives till age (x+n-1) but dies in the age period (x+n-1, x+n).

= P_{rob}(a person aged x survives for (n-1) years) times the P_{rob}(a person aged x+n-1 dies within one year.

Thus, by the compound probability theorem,

$$n/q_x = \frac{\ell_{x+n-1}}{\ell_x} \cdot \frac{d_{x+n-1}}{\ell_{x+n-1}}$$

$$= \frac{d_{x+n-1}}{\ell_x}$$

$$= \frac{\ell_{x+n-1} - \ell_{x+n}}{\ell_x} \qquad (2.10)$$

The probability that a person aged x will die between ages x+n and x+n+m is denoted by

$$n|m\, q_x = \frac{\ell_{x+n} - \ell_{x+n+m}}{\ell_x} \qquad (2.11)$$

From (2.2),

$$d_x = \ell_x - \ell_{x+1}$$

Therefore

$$\sum_{i=x}^{\omega} d_i = d_x + d_{x+1} + d_{x+2} + \dots + d_{\omega-1} + d_\omega$$

$$= (\ell_x - \ell_{x+1}) + (\ell_{x+1} - \ell_{x+2}) + (\ell_{x+2} - \ell_{x+3})$$

$$+ \dots + (\ell_{\omega-1} - \ell_\omega) + (\ell_\omega - \ell_{\omega+1})$$

If ω is the last age at which ℓ_x vanishes

i.e.,

$$\ell_\omega = o$$

Then

$$\ell_{\omega+1} = \ell_{\omega+2} = \dots = o$$

Therefore

$$\sum_{i=x}^{\omega} d_i = \sum_{i=1}^{\omega-1} d_i$$

$$= (\ell_x - \ell_{x+1}) + (\ell_{x+1} - \ell_{x+2}) + \dots + (\ell_{\omega-1} - \ell_\omega)$$

$$= \ell_x$$

Thus

$$\sum_{i=x}^{\omega-1} d_i = \ell_x \qquad (2.12)$$

2.2.2 Person - years lived

The number of person-years that ℓ_x persons, aged x are expected to live through (x, x+1) is

$$L_x = \int_x^{x+1} \ell_y \, dy$$

$$= \int_o^1 \ell_{x+t} \, dt \qquad (2.13)$$

If deaths are assumed to be uniformly distributed over the whole year or equivalently, if we assume the linearity of ℓ_{x+t} for $t \in (0, 1)$, then we get

$$L_x = \int_o^1 \ell_{x+t} \, dt$$

and

$$\ell_{x+t} = \ell_x - t\, d_x \qquad (2.14)$$

Thus

$$L_x = \int_o^1 (\ell_x - t\, d_x)dt$$

$$= \ell_x - \frac{1}{2} d_x$$

$$= \ell_x - \frac{1}{2}(\ell_{x+1} - \ell_n)$$

$$= \frac{1}{2}(\ell_x + \ell_{x+1}) \qquad (2.15a)$$

$$= \ell_{x+\frac{1}{2}} \qquad (2.15b)$$

by applying formula (2.14) and putting $t = \frac{1}{2}$. Using the finite difference technique, let

$$\Delta\ell_x = \ell_{x+1} - \ell_x \qquad (2.16)$$

Therefore

$$(1 + \Delta)\ell_x = \ell_x + \Delta\ell_x$$

$$= \ell_x + (\ell_{x+1} - \ell_x)$$

$$= \ell_{x+1} \qquad (2.17)$$

$$(1 + \Delta)^2 \ell_x = (1 + \Delta)(1 + \Delta)\ell_x$$

$$= (1 + \Delta)\ell_{x+1}$$

$$= \ell_{x+2} \qquad (2.18)$$

In general,

$$\ell_{x+t} = (1 + \Delta)^t \ell_x \qquad (2.19)$$

$$= \ell_x + t\,\Delta\,\ell_x, \text{ to first difference}$$

$$= \ell_x + t(\ell_{x+1} - \ell_x)$$

$$= \ell_x - t(\ell_x - \ell_{x+1})$$

$$= \ell_x - t\,d_x$$

Therefore

$$L_x = \int_0^1 \ell_{x+t}\,dt$$

$$= \int_0^1 (\ell_x - t\,\Delta\ell_x)dt$$

$$= t\,\ell_x - \frac{1}{2}t^2\,d_x$$

$$= \ell_x - \frac{1}{2}(\ell_x - \ell_{x+1})$$

$$= \frac{1}{2}(\ell_x + \ell_{x+1})$$

as in (2.15a).

The number of person-years that ℓ_x persons, aged x are expected to live through (x, x+n) is

$$_nL_x = \int_x^{x+n} \ell_y \, dy$$

$$= \int_o^n \ell_{x+t} \, dt \qquad (2.20)$$

An approximation to $_nL_x$ based an numerical quadrature is given by the formula

$$_nL_x \simeq \frac{n}{2}(\ell_x + \ell_{x+n}) + \frac{n}{24}(_nd_{x+n} - {}_nd_{x-n}) \qquad (2.21)$$

2.2.3 Total number of years lived T_x

The expected total number of years lived beyond exact age x by ℓ_x persons alive at that age is

$$T_x = L_x + L_{x+1} + L_{x+2} + \dots + L_{\omega-1} + L_\omega$$

where ω is the highest age attained.

But

$$L_\omega = \ell_\omega + \ell_{\omega+1} + \ell_{\omega+2} + \dots$$

$$= 0$$

since

$$\ell_\omega = \ell_{\omega+1} = \dots = 0$$

Therefore

$$T_x = L_x + L_{x+1} + \dots + L_{\omega-1} \qquad (2.22a)$$

$$= \sum_{i=o}^{\omega-x-1} L_{x+i} \qquad (2.22b)$$

or simply

$$T_x = \sum_{i=o}^{\infty} L_{x+i} \qquad (2.22c)$$

In the continuous case,

$$T_x = \int_{y=x}^{\omega} \ell_y \, dy$$

$$= \int_{o}^{\omega-x} \ell_{x+t} \, dt$$

$$= \int_{o}^{\infty} \ell_{x+t} \, d_t \qquad (2.23)$$

For an n-year age interval,

$$T_x = {}_nL_x + {}_nL_{x+n} + {}_nL_{x+2n} + \ldots$$

$$= \sum_{h=o}^{\infty} {}_nL_{x+nh} \qquad (2.24)$$

for fixed $n = 1, 2, 3, 4, 5 \ldots$

The expression in (2.22) can be re-written as follows:-

$$T_x = L_x + (L_{n+1} + L_{x+2} + \ldots)$$

$$= L_x + T_{x+1} \qquad (2.25)$$

Also (2.24) can be written as

$$T_x = {}_nL_x + \sum_{h=1}^{\infty} {}_nL_{x+nh}$$

$$= {}_nL_x + T_{x+1} \qquad (2.26)$$

2.2.4 The Force of Mortality

This is the instanteneous death rate at exactly age x, denoted by $\mu(x)$.

Thus

$$\mu(x) = \lim_{\Delta x \to o} \frac{\ell(x) - \ell(x+\Delta x)}{\ell(x)\Delta x}$$

$$= - \frac{1}{\ell(x)} \lim_{\Delta x \to o} \frac{\ell(x+\Delta x) - \ell(x)}{\Delta x}$$

$$= - \frac{1}{\ell(x)} \frac{d\ell}{dx} \qquad (2.27a)$$

$$= - \frac{d}{dx} \log \ell_x \qquad (2.27b)$$

To express $\ell(x)$ in terms of $\mu(x)$, we integrate formula (2.27b) and get

$$\int_o^y d \log \ell(x) = \int_o^y - \mu(x)dx$$

i.e.,

$$\log \frac{\ell_{(y)}}{\ell_{(o)}} = - \int_o^y \mu(x)dx$$

which implies

$$\frac{\ell_{(y)}}{\ell_{(o)}} = e^{- \int_o^y \mu(x)dx}$$

i.e.,

$$p(y) = e^{- \int_o^y \mu(x)dx} \qquad (2.28)$$

or

$$\ell_{(y)} = \ell_{(o)} \; e^{- \int_o^y \mu(x)dx} \qquad (2.29)$$

Since

$$\mu_x = - \frac{1}{\ell_x} \frac{d\ell_x}{dx} \qquad \text{cf } 2.27a$$

it follows that

$$- d\ell_x = \ell_x \, \mu_x \, dx \qquad (2.30)$$

which is the number of deaths occurring at the moment of attaining age x (out of ℓ_x persons alive at that age)

$$\mu_x \, dx = - \frac{d\ell_x}{\ell_x} \qquad (2.31)$$

represents the probability that a person of exact age x will die at that moment. Therefore $\ell_{x+t}\,\mu_{x+t}\,dt$ represents the number of deaths occurring at that moment of age x+t.
Since d_x is the number of deaths occurring between ages x and x+1, it follows that

$$d_x = \int_0^1 \ell_{x+t}\,\mu_{x+t}\,dt \qquad (2.32)$$

and

$$q_x = \frac{1}{\ell_x}\int_0^1 \ell_{x+t}\,\mu_{x+t}\,dt \qquad (2.33a)$$

$$= \int_0^1 {}_tP_x\,\mu_{x+t}\,dt \qquad (2.33b)$$

We should note that ${}_tP_x$ is the probability of a person aged x surviving up to age x+t. The probability that having reached age x+t, a person will die at that moment is $\mu_{x+t}\,dt$. Thus the probability that a person aged x will die at moment of age x+t is ${}_tP_x\,\mu_{x+t}\,dt$. Integrating this expression within the limits t = 0 and t = 1, the result is the probability that (x) will die within one year.

That is

$$q_x = \int_0^1 {}_tP_x\,\mu_{x+t}\,dt$$

as before (cf 2.33b).

If the function ${}_tP_x\,\mu_{x+t}\,dt$ is integrated between the limits 0 and ∞, the result is the total probability that (x) will die, which is a certainty.

Thus

$$Q_x = \int_0^\infty {}_tP_x\,\mu_{x+t}\,dt = 1 \qquad (2.34)$$

Beyond the limit age there are no survivors and the value of ${}_tP_x$, where $x+t > \omega$, is zero.

Therefore

$$Q_x = \int_0^\infty {}_tP_x\,\mu_{x+t}\,dt$$

$$= \int_0^{\omega-x} {}_tP_x\,\mu_{x+t}\,dt$$

If the limits be taken as t = n to t = n + 1, the result is

$$q_{x+n} = {}_{n/1}q_x$$

$$= \int_n^{n+1} {}_tp_x \, \mu_{x+t} \, dt$$

$$= -\frac{1}{\ell_x} \int_n^{n+1} \frac{d\,\ell_{x+t}}{dt} \, dt$$

$$= \frac{\ell_{x+n} - \ell_{x+n-1}}{\ell_x}$$

$$= \frac{d_{x+n}}{\ell_x} \qquad (2.35)$$

If the integration is from t = n to t = n+m then, we have

$$ {}_{n/m}q_x = {}_mq_{x+n} = \int_n^{n+m} {}_tp_x \, \mu_{x+t} \, dt$$

$$= \frac{\ell_{x+n} - \ell_{x+n+m}}{\ell_x} \qquad (2.36)$$

Next, from

$$\mu_x = -\frac{d\,\ell_x}{\ell_x}$$

or

$$\ell_x \mu_x = -\,d\,\ell_x$$

we can deduce that

$$\ell_{x+t}\,\mu_{x+t} = -\,d\,\ell_{x+t}$$

which implies that

$$\int_0^n \ell_{x+t}\,\mu_{x+t}\,dt = \int_0^n -\,d\,\ell_{x+t}$$

$$= \ell_x - \ell_{x+n}$$

Therefore

$$\frac{1}{\ell_x}\int_0^n \ell_{x+t}\,\mu_{x+t}\,dt = \frac{\ell_x - \ell_{x+n}}{\ell_x} = {}_nq_x.$$

By the stochastic approach Chiang' (1968), used the notion of force of mortality to determine the probability of survival as follows:-

Let

$\mu(x)\ \Delta x + o(\Delta x)$ = Probability of dying between age x and $x+\Delta x$ (2.37)

$P_{rob}(X \leqslant x)$ = $F(x)$

= Probability of dying at or before age x (2.38)

So

$F(x+\Delta x)$ = Probability of dying at or before age $x+\Delta x$

= Probability of dying at or before age x or probability of living up to age x and dying between ages $(x, x+\Delta x)$

$$= F(x) + [1 - F(x)]\,.[\mu(x)\Delta x + O(\Delta x)] \quad (2.39$$

Therefore

$$\lim_{\Delta x \to o} \frac{F(x+\Delta x) - F(x)}{\Delta x} = F'(x) = [1 - F(x)]\,\mu(x)$$

i.e.,

$$\frac{dF}{dx} = [1 - F(x)]\,\mu(x) \quad (2.40)$$

Using the dummy variable t,

$$\frac{dF}{dt} = [1-F(t)]\,\mu(t) \quad (2.41)$$

Therefore

$$\frac{dF/dt}{1 - F(t)} = \mu(t)$$

This implies

$$\int_o^x \frac{F'(t)}{1-F(t)}\, dt = \int_o^x \mu(t)\, dt$$

i.e.,

$$-\{\ln[1 - F(t)]\}_o^x = \int_o^x \mu(t)\, dt$$

i.e.,

$$-\ln[1 - F(x)] + \ln[1 - F(o)] = \int_o^x \mu(t)\, dt$$

But

$$F(0) = o$$

and

$$\ln 1 = 0$$

Therefore

$$-\ln[1 - F(x)] = \int_o^x \mu(t)\, dt$$

i.e.,

$$1 - F(x) = e^{-\int_o^x \mu(t)\, dt} \qquad (2.42)$$

which is the probability of living up to age x from birth.

2.2.5 Estimation of force of mortality μ_x

Various formulae can be obtained for the approximate value of μ_x. Assuming ℓ_x to be a function of the fourth degree, the value of μ_x can be expressed as follows:

Let

$$\ell_x = a + b\,x + cx^2 + d\,x^3 + e\,x^4 \qquad (2.43)$$

where a, b, c, d and e are constants.

Then

$$\frac{d\,\ell_x}{dx} = b + 2cx + 3dx^2 + 4ex^3$$

when $x = 0$,

$$\frac{d\,\ell_o}{dx} = b$$

Also

$$\ell_{-1} = a - b + c - d + e$$

$$\ell_{+1} = a + b + c + d + e$$

Therefore

$$\ell_{-1} - \ell_{+1} = -2b - 2d$$

Next

$$\ell_{-2} = a - 2b + 4c - 8d + 16e$$

and

$$\ell_{+2} = a + 2b + 4c + 8d + 16e$$

Therefore

$$\ell_{-2} - \ell_{+2} = -4b - 16d$$

Therefore

$$8(\ell_{-1} - \ell_{+1}) - (\ell_{-2} - \ell_{+2}) = -12b$$

Whence

$$\mu_o = -\frac{1}{\ell_o}\frac{d\,\ell_o}{dx}$$

$$= -\frac{1}{\ell_o}\,.\,b$$

$$= -\frac{12b}{12\ell_o}$$

$$= \frac{8(\ell_{-1} - \ell_{+1}) - (\ell_{-2} - \ell_{+2})}{12\ell_o} \qquad (2.42)$$

$$= \frac{8(\ell_{-1} - \ell_o + \ell_o - \ell_{+1}) - (\ell_{-2} - \ell_{-1} + \ell_{-1} - \ell}{12\ell_o}$$

i.e.,

$$\mu_o = \frac{8(d_{-1} + d_o) - (d_{-2} + \ell_{-1} - \ell_{+2})}{12\ell_o}$$

$$= \frac{7(d_{-1} + d_o) + (d_{-1}+d_o) - (d_{-2}+\ell_{-1} - \ell_{+2})}{12\ell_o}$$

$$= \frac{7(d_{-1}+d_o) + (\ell_{-1}-\ell_o) + (\ell_o-\ell_1) - (d_{-2}+\ell_{-1}-\ell_{+2})}{12\ell_o}$$

$$= \frac{7(d_{-1}+d_o) + (\ell_{-1}-\ell_1) - (d_{-2}+\ell_{-1} - \ell_{+2})}{12\ell_o}$$

$$= \frac{7(d_{-1} + d_o) - \ell_1 - d_{-2} + \ell_{+2}}{12\ell_o}$$

$$= \frac{7(d_{-1} + d_o) - (d_{-2} + \ell_1 - \ell_{+2})}{12\ell_o}$$

$$= \frac{7(d_{-1} + d_o) - (d_{-2} + d_{+1})}{12\ell_o} \qquad (2.43)$$

Therefore also,

$$\mu_x = \frac{8(\ell_{x-1} - \ell_{x+1}) - (\ell_{x-2} - \ell_{x+2})}{12\ell_x} \qquad (2.44)$$

$$= \frac{7(d_{x-1} + d_x) - (d_{x-2} + d_{x+1})}{12\ell_x} \qquad (2.45)$$

Using Taylor's expansion on ℓ_x, then

$$\ell_{x+h} = \ell_x + h\ell_x' + \frac{h^2}{2!}\ell_x'' + \frac{h^3}{3!}\ell_x''' + \ldots \qquad (2.46)$$

and

$$\ell_{x-h} = \ell_x - h\ell_x' + \frac{h^2}{2!}\ell_x'' - \frac{h^3}{3!}\ell_x''' + \ldots$$

Therefore

$$\ell_{x+h}-\ell_{x-h} = 2h\ell_x' + \frac{h^3}{3}\ell_x''' + \frac{h^5}{60}\ell_x^{v} + \ldots \qquad (2.47)$$

Assuming that ℓ_x''' and higher order differential coefficient are negligible, on putting h = 1, then

$$\ell_{x+1} - \ell_{x-1} = 2\ell_x' \qquad (2.48)$$

This implies

$$\mu_x = -\frac{\ell_x'}{\ell_x}$$

$$= \frac{\ell_{x-1} - \ell_{x+1}}{2\,\ell_x}$$

$$= \frac{(\ell_{x-1} - \ell_x) + (\ell_x - \ell_{x+1})}{2\,\ell_x} \qquad (2.49)$$

$$= \frac{d_{x-1} + d_x}{2\,\ell_x}, \quad x > 1 \qquad (2.50)$$

A better approximation to μ_x is obtained on retaining terms up to fourth order differential coefficient of ℓ_x and neglecting higher order differential coefficients.

Thus on putting h = 1 and h = 2 respectively in (2.47), we get

$$\ell_{x+1} - \ell_{x-1} = 2\,\ell_x' + \frac{1}{3}\ell_x''' \qquad (2.51)$$

and

$$\ell_{x+2} - \ell_{x-2} = 4\ell_x' + \frac{8}{3}\ell_x''' \qquad (2.52)$$

Eliminating ℓ_x''' between these equations, multiply (2.51) by 8 and subtract the result from (2.52) i.e.,

$$8(\ell_{x+1} - \ell_{x-1}) = 16\,\ell_x' + \frac{8}{3}\ell_x'''$$

$$\ell_{x+2} - \ell_{x-2} = 4\,\ell_x' + \frac{8}{3}\ell_x'''$$

$$8(\ell_{x+1} - \ell_{x-1}) - (\ell_{x+2} - \ell_{x-2}) = 12\,\ell_x' \qquad (2.53)$$

Therefore

$$\mu_x = - \frac{\ell'_x}{\ell_x} = - \frac{12\ell'_x}{12\ell_x}$$

$$= \frac{8(\ell_{x-1} - \ell_{x+1}) - (\ell_{x-2} - \ell_{x-2})}{12\ \ell_x}$$

as in (2.44).

Other approximate expressions can be obtained from the relation connecting the differential operator with the finite difference operator Δ:

$$D = \log(1 + \Delta) = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \ . \ . \ . \qquad (2.54)$$

Thus

$$\mu_x = - \frac{1}{\ell_x} \frac{d\ell_x}{dx}$$

$$= - \frac{1}{\ell_x} D\ell_x$$

$$= - \frac{1}{\ell_x} (\Delta\ell_x - \frac{1}{2} \Delta^2 \ell_x + \frac{1}{3} \Delta^3 \ell_x - \ . \ . \ .) \qquad (2.55)$$

$$= \frac{1}{\ell_x} (d_x - \frac{1}{2} \Delta\, d_x + \frac{1}{3} \Delta^2 d_x - \ . \ . \ . \ ) \qquad (2.56)$$

Alternatively

$$\mu_x = - \frac{d}{dx} \log \ell_x$$

$$= - (\Delta \log \ell_x - \frac{1}{2} \Delta^2 \log \ell_x + \frac{1}{3} \Delta^3 \log \ell_x - \ . \ . \ . \ )$$

$$= \text{colog } p_x - \frac{1}{2} \Delta \text{ colog } p_x + \frac{1}{3} \Delta^2 \text{ colog } p_x - \ . \ . \ .$$

(2.57)

We have shown in (2.29), that

$$\ell_x = \ell_o \, e^{-\int_o^x \mu_y \, dy}$$

Therefore

$$\ell_{x+n} = \ell_o \, e^{-\int_o^{x+n} \mu_y \, dy}$$

Then

$$_nP_x = \frac{\ell_{x+n}}{\ell_x}$$

$$= \frac{\ell_o \, e^{-\int_o^{x+n} \mu_y \, dy}}{\ell_o \, e^{-\int_o^x \mu_y \, dy}}$$

$$= e^{-\int_x^{x+n} \mu_y \, dy} \qquad (2.58)$$

Let

$$y = x + t$$

This implies that

$$t = y - x$$

and

$$dt = dy$$

Therefore when $y = x, \quad t = o$

and when $y = x+n, \quad t = n$

Therefore

$$_nP_x = e^{-\int_o^n \mu_{x+t} \, dt} \qquad (2.59)$$

Now, when $n = 1$,

$$_1P_x = P_x = p(x) = e^{-\int_o^1 \mu_{x+t} \, dt} \qquad (2.60)$$

Taking logarithms, we have

$$\log p(x) = - \int_o^1 \mu_{x+t} \, dt$$

The definite integral represents the mean value of μ between the ages x and x + 1. If we approximate this mean value by $\mu_{x+\frac{1}{2}}$, we have

$$\mu_{x+\frac{1}{2}} \simeq - \log p_x$$

If we integrate between t = - 1 and t = 1, we find

$$\int_{-1}^{1} \mu_{x+t}\, dt = - \log p_{x-1} - \log p_x \qquad (2.61)$$

and this is twice the mean value of μ between the ages x - 1 and x+1. This leads to the following approximation:

$$\mu_x \simeq - \frac{1}{2} (\log p_{x-1} + \log p_x) \qquad (2.62)$$

$$= \frac{1}{2} (\log \ell_{x-1} - \log \ell_{x+1}) \qquad (2.63)$$

2.2.6 Laws of Mortality

We now wish to express the relation (2.29) for various types of force of mortality $\mu(x)$. Abraham De Moivre (1725) proposed a very simple law of mortality, namely

$$\mu(x) = \frac{1}{\omega - x} \qquad (2.64)$$

Therefore

$$\ell(y) = \ell(o)\, e^{-\int_0^y \mu(x)dx} \qquad (2.29)$$

becomes

$$\ell(y) = \ell(o)\, e^{-\int_0^y \frac{1}{\omega-x} dx}$$

$$= \ell(o)\, e^{\log \omega - x \Big|_0^y}$$

$$= \ell(o)\, e^{\log \frac{\omega-y}{\omega}}$$

i.e.,

$$\ell(y) = \frac{\ell(o)}{\omega}(\omega-y) \qquad (2.65a)$$

$$= k(\omega-y) \qquad (2.65b)$$

where k is a constant and ω is the highest attainable age.

The most famous mathematical expression of the force of mortality is the Gompertz-Makeham formula. In a paper on the "Law of Human Mortality Benjamin Gompertz (1825) attributed death to either of two causes:

1. due to chance
2. deterioration of the power to withstand destruction.

In deriving his law of mortality, however, he considered only the second cause and stipulated that "man's power to resist death decreases at a rate proportional to the power itself". Since $\mu(x)$ is a measure of man's susceptibility to death, Gompertz used

$$R(x) = \frac{1}{\mu(x)} \qquad (2.66)$$

as a measure of man's resistance to death. He then translated his postulation into the differential equation

$$\frac{d}{dx} R(x) = - hR(x) \qquad (2.67)$$

where h is a positive constant.
Solving (2.67) we get

$$\int \frac{1}{R(x)}\frac{d}{dx} R(x)\, dx = \int - h\, dx$$

which implies

$$\log R(x) = - hx + k$$

i.e.,

$$R(x) = e^{-hx + k}$$

i.e.,

$$\frac{1}{\mu(x)} = e^{-hx + k}$$

Therefore

$$\mu(x) = e^{hx - k}$$

$$= e^{-k} e^{hx}$$

$$= B\,C^{x} \qquad (2.68)$$

where B and C are parameters (constants).

So

$$\ell(y) = \ell(o)\, e^{-\int_o^y BC^x\,dx}$$

$$= \ell(o)\, e^{-B\int_o^y C^x\,dx} \qquad (2.69)$$

Let

$$Z = C^x$$

then

$$\log Z = x \log C$$

Thus

$$\frac{d}{dx} \log Z = \log C$$

i.e.,

$$\frac{1}{Z}\frac{dz}{dx} = \log C$$

Therefore

$$dx = \frac{dz}{z \log C}$$

Therefore, from (2.69),

$$\int_o^y C^x\,dx = \int_{z=1}^{C^y} Z \,.\, \frac{d\,z}{z \log C}$$

$$= \int_1^{C^y} \frac{d\,z}{\log C}$$

$$= \frac{C^y - 1}{\log C}$$

Therefore (2.69) becomes

$$\ell(y) = \ell(o)\, e^{-B\left(\frac{C^y - 1}{\log C}\right)}$$

$$= \ell(o)\, e^{\left[-\frac{B\,C^y}{\log C} + \frac{B}{\log C}\right]}$$

$$= \ell(o) \exp\left(-\frac{B}{\log C}\right)C^y \exp\frac{B}{\log C} \qquad (2.70)$$

Let

$$g = \exp - \frac{B}{\log C} \qquad (2.71)$$

and

$$k = \ell(o) \exp\frac{B}{\log C} \qquad (2.72)$$

Then (2.69) becomes

$$\ell(y) = k\, g^{C^y} \qquad (2.73)$$

Makeham (1860) suggested the modification

$$\mu(x) = A + BC^x \qquad (2.74)$$

to restore the missing component "chance" to the Gompertz formula.

So

$$\ell(y) = \ell(o)\, e^{-\int_o^y (A+BC^x)dx}$$

$$= \ell(o)\, e^{-\int_o^y Adx}\, e^{-\int_o^y BC^x dx}$$

$$= \ell(o)\, e^{Ay}\, e^{-\int_o^y BC^x dx}$$

$$= e^{-Ay}\, \ell(o)\, e^{-\int_o^y BC^x dx} \qquad (2.75)$$

Using (2.69) and (2.73) and letting

$$S = e^{-A},$$

then (2.75) becomes

$$\ell(y) = k\, s^y\, g^{c^y}, \qquad (2.76)$$

For

$$\mu(x) = A + Hx + BC^x \quad (2.77)$$

$$\begin{aligned} \ell(y) &= \ell(o) \exp[- \int_o^y (A + Hx + BC^x) dx] \\ &= \ell(o) \exp[- (Ay + \frac{H}{2} y^2 + \int_o^y BC^x dx] \\ &= \ell(o) \exp[- \int_o^y BC^x dx] \exp[- Ay - \frac{H}{2} y^2] \\ &= k g^{C^y} e^{-Ay} e^{-\frac{H}{2} y^2} \\ &= k g^{C^y} s^y u^{y^2} \\ &= k s^y u^{y^2} g^{c^y} \end{aligned} \quad (2.78)$$

where

$$u = e^{-H/2} \quad (2.79)$$

and the other parameters are as have been defined before.

In 1932, in the Journal of the Institute of Actuaries, Perk proposed to modify the Gompertz - Makeham formula to

$$\mu(x) = \frac{1 + BC^x}{1 + DC^x} \quad (2.80)$$

Remark: Gompertz - type laws are primarily fitting with adult ages and not for infant and child mortality.

About 1870 Oppermann suggested a formula for graduation of infant and childhood mortality. He defined

$$\mu(x) = \frac{a}{\sqrt{x}} + b + c\sqrt{x} \quad (2.81)$$

or in terms of the continuous survivorship function

$$\log \ell(x) = A\sqrt{x} + Bx + Cx^{3/2} \quad (2.82)$$

where log could be either the natural logarithmic function or $\log_{10}$

Using the method of unweighted least squares techniques, we have the following normal equations

$$A\sum_i x_i + B\sum_i x_i^{3/2} + C\sum_i x_i^2 = \sum_i \sqrt{x_i}\,\log \ell(x_i) \qquad (2.83\text{ i})$$

$$A\sum_i x_i^{3/2} + B\sum_i x_i^{3/2} + C\sum_i x_i^{5/2} = \sum_i x_i \log \ell(x_i) \qquad (2.83\text{ii})$$

$$A\sum_i x_i^2 + B\sum_i x_i^{5/2} + C\sum_i x_i^3 = \sum_i x_i^{3/2} \log \ell(x_i) \qquad (2.83\text{iii})$$

Thiele (1872) was of the opinion that such formulae should take into account the differences in mortality behaviour during the major epochs of life. Thus he wanted to partition the survivorship curve into three components.

For childhood he used the formula

$$\mu_1(x) = a_1 \exp(-b_1 x) \qquad (2.84)$$

For adult ages,

$$\mu_2(x) = a_2 \exp(-\frac{1}{2} b_2^2 (x-c)^2) \qquad (2.85)$$

and for old ages, the formula is

$$\mu_3(x) = a_3 \exp(b_3 x) \qquad (2.86)$$

The formula meant for graduation of mortality throughout all ages was written

$$\mu(x) = \mu_1(x) + \mu_2(x) + \mu_3(x) \qquad (2.87)$$

For studies of life span of materials, Weibull (1939) recommended

$$\mu(x) = \mu a x^{a-1} \qquad (2.88)$$

where μ and a are constants.

Therefore the survival function

$$\ell(y) = \ell(o)\, e^{-\int_o^y \mu a\, x^{a-1} dx}$$

$$= \ell(o)\, e^{-\mu y^a} \qquad (2.89)$$

Weibull distribution is extensively used in reliability theory.

If

$$\mu(x) = \mu \text{ , constant} \tag{2.90}$$

then

$$\ell(y) = \ell(o)\, e^{-\int_o^y \mu\, dx}$$

$$= \ell(o)\, e^{-\mu y} \tag{2.91}$$

which plays a central role in the problem of life testing.

According to the Landahl model,

$$\mu(x) = \frac{p}{1+kx} \tag{2.92}$$

where p and k are parameters representing the combined effects of all risks which may result in death.

Therefore

$$\ell(y) = \ell(o)\, e^{-\int_o^y \frac{p}{1+kx}\, dx}$$

$$= \ell(o)\, e^{-{}^p/_k \log(1+k\,y)}$$

$$= \ell(o) \cdot \frac{1}{(1+k\,y)^{p/k}} \tag{2.93}$$

2.2.7 Expectation of Life

We know that the total probability that a person aged x will die is a certainty.

That is

$$Q_x = \int_o^\infty {}_t p_x\, \mu_{x+t}\, dt = 1$$

as shown in formula (2.34).

The function ${}_t p_x\, \mu_{x+t}\, dt$ is the probability that a person aged x will die at moment of age x+t. This function is a probability density. Therefore its expectation is given by

$$e^o_x = \int_0^\infty t \; {}_tp_x \; \mu_{x+t} dt \qquad (2.94)$$

$$= \int_0^\infty t \cdot \frac{\ell_{x+t}}{\ell_x} \cdot \mu_{x+t} \, dt$$

$$= \frac{1}{\ell_x} \int_0^\infty t \cdot \ell_{x+t} \cdot \mu_{x+t} \, dt$$

$$= \frac{1}{\ell_x} \int_0^\infty t \cdot \ell_{x+t} \left[- \frac{1}{\ell_{x+t}} \frac{d\,\ell_{x+t}}{dt} \right] dt$$

$$= - \frac{1}{\ell_x} \int_0^\infty t \frac{d\,\ell_{x+t}}{dt} \cdot dt \qquad (2.95)$$

Integrating by parts, let

$v = t$ and $du = d\ell_{x+t}$

Therefore

$dv = dt$ and $u = \ell_{x+t}$

Therefore (2.95) becomes

$$e^o_x = - \frac{1}{\ell_x} \; t \cdot \ell_{x+t} \Big|_0^\infty - \int_0^\infty \ell_{x+t} \, dt$$

$$= - \frac{1}{\ell_x} \cdot 0 - \int_0^\infty \ell_{x+t} \, dt$$

$$= \frac{1}{\ell_x} \int_0^\infty \ell_{x+t} \, dt$$

$$= \frac{T_x}{\ell_x} \qquad (2.96)$$

which is the required relation.

2.2.8 Age Specific Mortality Rate

This is the mortality rate for the specific age group. It is the ratio of the deaths recorded during a year to the mid-year population. Thus the observed death rate ${}_nM_x$ is defined by

$$ {}_nM_x = \frac{{}_nD_x}{{}_nK_x} \qquad (2.97)$$

where ${}_nD_x$ is the number of deaths of people aged between x and x+n and ${}_nK_x$ is the mid-population for that age group. The age specific mortality rate at exact age x is defined as the force of mortality denoted by $\mu(x)$.

That is

$$\mu(x) = \lim_{n \to o} {}_nM_x \qquad (2.98)$$

Using finite approximations, we can express the survival function $\ell(a)$ by

$$\ell(a) = \ell(o)\, e^{-\int_o^a \mu(x)\, dx}$$

$$= \ell(o)\, e^{-{}_1M_o - {}_1M_1 - {}_1M_2 - \cdots - {}_1M_{a-1}} \qquad (2.99)$$

assuming linearity in one year age interval.

The life table central death rate is denoted by small m. That is

$${}_nm_x = \frac{\ell_x - \ell_{x+n}}{{}_nL_x}$$

$$= \frac{{}_nd_x}{{}_nL_x} \qquad (2.100)$$

In particular

$$m_x = \frac{d_x}{L_x} \qquad (2.101)$$

which is the number of deaths per person-years lived in x, x+1 .

2.2.9 <u>The relationships between ${}_nq_x$ and ${}_nm_x$</u>

<u>Case 1:</u> Using the notion of fraction of last years of life.

Each of the ${}_nd_x$ people who die during the interval (x, x+n) has lived x complete-years plus some-fraction say 'a'. So

$$a = \frac{\text{Person-years lived by those who die in } (x, x+n)}{\text{Number who die in } (x, x+n)}$$

$$= \frac{{}_nL_x - n \, . \, \ell_{x+n}}{\ell_x - \ell_{x+n}}$$

This implies

$$a(\ell_x - \ell_{x+n}) = {}_nL_x - n \, . \, \ell_{x+n}$$

i.e.,

$${}_nL_x = a \, \ell_x + (n - a) \, \ell_{x+n} \qquad (2.102)$$

From the definitions

$${}_nm_x = \frac{\ell_x - \ell_{x+n}}{{}_nL_x}$$

and

$${}_nq_x = \frac{\ell_x - \ell_{x+n}}{\ell_x}$$

we get

$${}_nq_x = {}_nm_x \, . \, \frac{{}_nL_x}{\ell_x} \qquad (2.103)$$

Therefore

$${}_nq_x = {}_nm_x \, . \, \frac{a \, \ell_x + (n-a) \, \ell_{x+n}}{\ell_x}$$

$$= [{}_nm_x \, [\, a + (n - a) \, {}_np_x]$$

$$= {}_nm_x \, [\, a + (n - a)(1 - {}_nq_x)]$$

$$= {}_nm_x \, (a + n - n \, . \, {}_nq_x - a + a \, . \, {}_nq_x)$$

$$= {}_nm_x \, [\, n - (n - a) \, {}_nq_x \,]$$

$$= n \, . \, {}_nm_x - {}_nm_x \, (n - a) \, {}_nq_x$$

i.e.,

$$1 + (n - a) \cdot {}_n m_x \cdot {}_n q_x = n \cdot {}_n m_x$$

Therefore

$$ {}_n q_x = \frac{n \cdot {}_n m_x}{1 + (n-a)\, {}_n m_x} \qquad (2.104)$$

When

$$a = \frac{n}{2} \qquad (2.105)$$

then

$$ {}_n q_x = \frac{n \cdot {}_n m_x}{1 + \frac{n}{2} \cdot {}_n m_x} \qquad (2.106)$$

and further when

$$n = 1 \qquad (2.107)$$

then

$$q_x = \frac{m_x}{1 + \frac{1}{2} m_x} \qquad (2.108a)$$

$$= \frac{2\, m_x}{2 + m_x} \qquad (2.108b)$$

Case 2: Relationship between ${}_n q_x$ and ${}_n m_x$ when the survival function is linear

Let

$$\ell_x = a + bx \qquad (2.109)$$

Therefore, by definition

$$ {}_n m_x = \frac{\ell_x - \ell_{x+n}}{{}_n L_x}$$

$$= \frac{\ell_x - \ell_{x+n}}{\int_x^{x+n} \ell_y \, dy}$$

$$= \frac{(a + bx) - (a + b(x + n))}{\int_x^{x+n} (a+by)\,dy}$$

$$= \frac{-bn}{(ay + \frac{b}{2}y^2)\Big|_x^{x+n}}$$

$$= \frac{-bn}{a(x+n) + \frac{b}{2}(x+n)^2 - ax - \frac{b}{2}x^2}$$

$$= \frac{-bn}{a(x+n) + \frac{b}{2}(x^2 + 2xn+n^2) - ax - \frac{b}{2}x^2}$$

$$= \frac{-bn}{an + \frac{b}{2}(2xn + n^2)}$$

$$= \frac{-b}{a + \frac{b}{2}(2x + n)}$$

i.e.,

$$_nm_x = \frac{-b}{a + bx + \frac{bn}{2}} \qquad (2.110)$$

Next

$$_nq_x = \frac{\ell_x - \ell_{x+n}}{\ell_x}$$

$$= \frac{-bn}{a + bx} \qquad (2.111)$$

But from (2.110)

$$a + bx = -\frac{b}{_nm_x} - \frac{bn}{2}$$

which when substituted in (2.111), we get

$$ {}_nq_x = \frac{-\,b\,n}{-\dfrac{b}{{}_nm_x} - \dfrac{b\,n}{2}} $$

$$ = \frac{n}{\dfrac{1}{{}_nm_x} + \dfrac{n}{2}} $$

$$ = \frac{2n \,.\, {}_nm_x}{2 + n \,.\, {}_nm_x} \qquad (2.112) $$

Case 3: <u>Relationship between ${}_nq_x$ and ${}_nm_x$ when the survival function is exponential</u>

Let

$$ \ell_x = e^{\,a+bx} \qquad (2.113) $$

Then

$$ {}_nm_x = \frac{e^{\,a+bx} - e^{\,a+b(x+n)}}{\int_x^{x+n} e^{a+by}\, dy} $$

$$ = \frac{e^{a+bx} - e^{a+b(x+n)}}{\dfrac{e^{a+by}}{b}\Big|_x^{x+n}} $$

$$ = \frac{b\left(e^{a+bx} - e^{a+b(x+n)}\right)}{e^{a+b(x+n)} - e^{a+bx}} $$

$$ = -\,b \qquad (2.114) $$

Next

$$ {}_nq_x = \frac{\ell_x - \ell_{x+n}}{\ell_x} $$

$$ = \frac{e^{a+bx} - e^{a+b(x+n)}}{e^{a+bx}} $$

$$ = 1 - e^{bn} \qquad (2.115) $$

$$= \frac{-\frac{\ell_x^2}{\ell_x^2} - \frac{d\ell_x}{dx}\int_x^\infty \ell_y\,dy}{\ell_x^2}$$

$$= \frac{-1 - \frac{1}{\ell_x}\frac{d\ell_x}{dx}\int_x^\infty \ell_y\,dy}{\ell_x}$$

i.e.,

$$\frac{d}{dx}\,\overset{o}{e}_x = -1 + \mu_x \frac{T_x}{\ell_x}$$

$$= -1 + \mu_x\,\overset{o}{e}_x$$

Therefore

$$\frac{d\,\overset{o}{e}_x}{dx} + 1 = \mu_x\,\overset{o}{e}_x \qquad (2.117)$$

<u>Case 2</u>: Relationship between the central death rate m_x and the force of mortality μ_x.

By definition,
the central death rate is

$$m_x = \frac{d_x}{L_x}$$

and the force of mortality is

$$\mu_x = \frac{1}{\ell_x}\frac{d\ell_x}{dx} = -\frac{d}{dx}\log \ell_x$$

Also

$$L_x = \int_o^1 \ell_{x+t}\,dt$$

Therefore

$$\frac{d\,L_x}{dx} = \int_0^1 \frac{d}{dx}\, \ell_{x+t}\;dt$$

i.e.,

$$\frac{d\,L_x}{dx} = \ell_{x+1} - \ell_x$$

$$= -(\ell_x - \ell_{x+1})$$

$$= -d_x \qquad (2.118)$$

Dividing both sides by L_x we get

$$\frac{1}{L_x}\;\frac{dL_x}{d_x} = -\frac{d_x}{L_x} = -m_x$$

Therefore

$$\frac{1}{L_x}\;\frac{dL_x}{dx} = -m_x \qquad (2.119)$$

or

$$m_x = -\frac{d}{dx}\log L_x \qquad (2.120)$$

Assuming linearity,

$$L_x = \ell_{x+\frac{1}{2}} \quad \text{cf} \quad (2.15a)$$

Then

$$m_x = -\frac{d}{dx}\log \ell_{x+\frac{1}{2}}$$

$$= \mu_{x+\frac{1}{2}} \qquad (2.121)$$

by definition of μ_x.

CHAPTER 3

THE STABLE POPULATION

3.1 Birth rate, Death rate and Age distribution in general:

Suppose n(a,t) is the number of persons at age a and time t. If N(t) is the total number of persons at time t, then the proportion of persons at age a and time t is

$$c(a,t) = \frac{n(a,t)}{N(t)} \tag{3.1}$$

Next, consider a female population whose annual death rate at age a and time t is denoted by $\mu(a,t)$. So the number of deaths at age a and time t is $n(a,t) \times \mu(a,t)$. The total number of deaths for all ages at time t is therefore

$$D(t) = \int_0^w n(a,t)\, \mu(a,t)\, da \tag{3.2}$$

Similarly, if the annual rate of bearing a female child at age a and time t is denoted by m(a,t), then the total number of births at time t is

$$B(t) = \int_0^w n(a,t)\, m(a,t)\, da \tag{3.3}$$

Formula (3.3) can be expressed in another way as follows:

Suppose p(a,t) is the probability of surviving from birth up to age a at time t. Then the number of persons at age a and time t must be the survivors of births (t-a) years ago.

Therefore

$$n(a,t) = B(t-a)\, p(a,t) \tag{3.4}$$

So formula (3.3) can be expressed as

$$B(t) = \int_0^w B(t-a)\, p(a,t)\, m(a,t)\, da$$

$$= \int_{\alpha}^{\beta} B(t-a)\ p(a,t)\ m(a,t)\ da \qquad (3.5)$$

where α and β are the lower and upper limits of child-bearing.

We now wish to determine crude death and birth rates using formulae (3.2) and (3.3). Crude death rate at time t is defined as

$$d(t) = \frac{D(t)}{N(t)}$$

$$= \int_{o}^{w} \frac{n(a,t)}{N(t)}\ \mu(a,t)\ da$$

$$= \int_{o}^{w} c(a,t)\ \mu(a,t)\ da \qquad (3.6)$$

Crude birth rate at time t is

$$b(t) = B(t)/N(t)$$

$$= \int_{o}^{w} \frac{n(a,t)}{N(t)}\ m(a,t)\ da$$

$$= \int_{o}^{w} c(a,t)\ m(a,t)\ da \qquad (3.7)$$

3.2 The concept of a stable Population:

Lopez (1967) showed that two populations with the same sequence of fertility and mortality schedules over a long period of time, but with different age distributions a long time ago, have the same current age distribution.

Alternatively, the age distribution of a closed population is determined by the history of its fertility and mortality in the recent past and does not depend on age distribution or fertility and mortality in the remote past.

Definition: A population that is established by a prolonged regime of unchanging fertility and mortality is called a stable population.

So, a stable population has a fixed age composition, constant birth and death rates and hence a constant rate of increase. In other words, age distribution, birth rate, death rate and the rate of increase are independent of time. So.

$$c(a,t) = c(a)$$

$$b(t) = b$$

and

$$d(t) = d$$

Thus the birth rate is

$$b = \int_0^w c(a)\, m(a)\, da \qquad (3.8)$$

and the death rate is

$$d = \int_0^w c(a)\, \mu(a)\, da \qquad (3.9)$$

With fixed age composition, birth and death rates are constant; and the difference between them, which in a closed population is the rate of increase, is also constant. That is

$$b - d = r \qquad (3.10)$$

The total number of persons is

$$N(t) = N(0)e^{rt} \qquad (3.11)$$

Because of constant birth and death rates, the annual number of births and deaths follow similar paths.
That is

$$B(t) = B(0)e^{rt} \qquad (3.12)$$

and

$$D(t) = D(0)e^{rt} \qquad (3.13)$$

3.3 Relations in Stable Population:

From (3.1) and (3.4)

$$c(a,t) = \frac{n(a,t)}{N(t)}$$

$$= \frac{B(t-a)\, p\,(a,t)}{N(t)}$$

$$= \frac{B(t-a)}{N(t-a)} \cdot \frac{N(t-a)}{N(t)}\; p(a,t)$$

But by definition of birth rate,

$$\frac{B(t-a)}{N(t-a)} = b(t-a)$$

Therefore

$$c(a,t) = b(t-a)\, \frac{N(t-a)}{N(t)}\, p(a,t) \tag{3.14}$$

If a population is growing at the rate 'r', then the population size at time t can be expressed in terms of the population size 'a' years ago as

$$N(t) = N(t-a)e^{ra} \tag{3.15a}$$

which implies

$$N(t-a) = N(t)e^{-ra} \tag{3.15b}$$

Substituting (3.15b) into (3.14) we get

$$c(a,t) = b(t-a)e^{-ra}\, p(a,t) \tag{3.16}$$

But in a stable population, mortality, fertility and age distribution are independent of time t. So (3.16) becomes

$$\boxed{c(a) = be^{-ra}\, p(a)} \tag{3.17}$$

This implies

$$\int_0^w c(a)da = \int_0^w be^{-ra}\, p(a)\, da$$

i.e.,

$$1 = b \int_0^w e^{-ra}\, p(a)\, da$$

which implies

$$\boxed{b = 1/\int_0^w e^{-ra}\, p(a)\, da} \qquad (3.18)$$

Using the expression (3.8) in section 3.2 for birth rate and formula (3.17) then,

$$b = \int_0^w c(a)\, m(a)\, da$$

$$= \int_0^w be^{-ra}\, p(a)\, m(a)\, da$$

i.e.,

$$\boxed{1 = \int_0^w e^{-ra}\, p(a)\, m(a)\, da} \qquad (3.19)$$

which is known as the characteristic equation of the stable population.

3.4 Stationary Population:

A special case of a stable population is when the rate of increase is zero. Then we have a zero growth population which is also known as a stationary population. In a stationary population, the total number of persons is the same every year. This is so because the number of births is the same so the number of deaths. A life table is a stationary population. So when

$$r = 0,$$

$$c(a) = b\, p(a) \qquad (3.20)$$

$$b = \frac{1}{\int_0^w p(a)da}$$

$$= \frac{1}{\frac{1}{\ell(o)} \int_o^w \ell(a)da}$$

$$= \frac{1}{\frac{1}{\ell(o)} T_o}$$

$$= \frac{1}{e_o^o} \qquad (3.21)$$

where

$\ell(a)$ = survivors at age a

$\ell(o)$ = the initial cohort

T_o = total population

e_o^o = expectation of life at birth

Next, since from formula (3.10),

$$b-d = r ,$$

then $$b = d \text{ (for } r = o) \qquad (3.22)$$

3.5 Five-year Age-intervals:

Let ${}_5C_a$ be the age distribution between exact ages 5 and a+5. Then

$${}_5C_a = \int_a^{a+5} c(x)dx$$

$$= \int_a^{a+5} be^{-rx} p(x)dx$$

$$= \int_a^{a+5} be^{-rx} \frac{\ell(x)}{\ell(o)} dx$$

$$= \frac{b}{\ell(o)} \int_a^{a+5} e^{-rx} \ell(x)dx$$

Now, let us replace x in e^{-rx} by the mid-value between a and a+5 which is a+2.5. Then

$$ {}_5C_a \simeq \frac{b}{\ell(o)} \int_a^{a+5} e^{-r(a+2.5)} \ell(x)dx $$

$$ = \frac{b}{\ell(o)} e^{-r(a+2.5)} \int_a^{a+5} \ell(x)dx $$

$$ = \frac{b}{\ell(o)} e^{-r(a+2.5)} {}_5L_a \qquad (3.23) $$

where ${}_5L_a$ is the person-years lived between ages a and a+5. Formula (3.23) implies that

$$ \int_o^w {}_5C_a \, da \simeq b \int_o^w e^{-r(a+2.5)} \frac{{}_5L_a}{\ell_o} da $$

But

$$ \int_o^w {}_5C_a \, da = 1 $$

Therefore

$$ b \simeq \frac{1}{\int_o^w e^{-r(a+2.5)} \frac{{}_5L_a}{\ell_o} da} \qquad (3.24) $$

Therefore

$$ {}_5C_a \simeq \frac{e^{-r(a+2.5)} {}_5L_a}{\int_o^w e^{-r(a+2.5)} {}_5L_a \, da} \qquad (3.25) $$

In the discrete form

$$ b \simeq \frac{1}{\sum_{a=o}^{w} e^{-r(a+2.5)} \frac{{}_5L_a}{\ell_o}} \qquad (3.26) $$

and

$$ {}_5C_a \simeq b\, e^{-r(a+2.5)} \frac{{}_5L_a}{\ell(o)} = \frac{e^{-r(a+2.5)} {}_5L_a}{\sum_{a=o}^{w} e^{-r(a+2.5)} {}_5L_a} . \qquad (3.27) $$

Next, given proportions of two five-year-age intervals, say ${}_5C_a$ and ${}_5C_y$, how do we determine the rate of increase and birth rate?

Consider

$$\frac{{}_5C_x}{{}_5C_y} \simeq \frac{e^{-r(x+2.5)}\ {}_5L_x}{\sum\limits_{x=o}^{w} e^{-r(x+2.5)}\ {}_5L_x} \cdot \frac{\sum\limits_{y=o}^{w} e^{-r(y+2.5)}\ {}_5L_y}{e^{-r(y+2.5)}\ {}_5L_y}$$

$$= e^{r(y-x)} \frac{{}_5L_x}{{}_5L_y}$$

which implies that

$$r \simeq \frac{1}{y-x} \ln \frac{{}_5C_x}{{}_5C_y} \cdot \frac{{}_5L_y}{{}_5L_x} \qquad (3.28)$$

Birth rate b can be obtained by substituting this value of r into the formula (3.24) or (3.26) where 'a' can be replaced by x or y.

3.6 Determining the Intrinsic rate of increase

The intrinsic rate of natural increase, is the growth rate attained by a population under a fixed regime of survivership and fertility; that is, it is the growth rate of a stable population.

A number of ways are available for calculating this rate, whose inter-relation throws light on the process by which a population grows.

Fundamentally, the problem of determining the intrinsic rate is trying to solve the characteristic equation of a stable population by numerical methods. Such solutions have been sought by Lotka (1925), Coale (1957), Pollard (1970), McCann (1973) and Keyfitz (1975).

We shall show the techniques used by Lotka and Coale in this text.

Lotka's Method

The characteristic equation is given by formula (3.19) as

$$\int_{o}^{w} e^{-ra}\, p(a)\, m(a)da = 1$$

which is equivalent to

$$\int_{\alpha}^{\beta} e^{-ra}\, p(a)\, m(a)da = 1$$

Net Reproductive Rate is defined by

$$NRR = \int_{\alpha}^{\beta} p(a)\, m(a)da$$

which implies

$$1 = \int_{\alpha}^{\beta} \frac{p(a)\, m(a)}{NRR}\, da \qquad (3.29)$$

So we can look at p(a) m(a)/NRR as a probability density function.

Define

$$R_n = \int_{\alpha}^{\beta} a^n\, p(a)\, m(a)da \qquad (3.30)$$

Then

$$R_o = \int_{\alpha}^{\beta} p(a)m(a)da = NRR$$

$$R_1 = \int_{\alpha}^{\beta} a\, p(a)\, m(a)da$$

and

$$R_2 = \int_{\alpha}^{\beta} a^2\, p(a)\, m(a)da$$

Divide the characteristic equation by R_o. So we have

$$\int_{\alpha}^{\beta} \frac{e^{-ra}\, p(a)m(a)da}{R_o} = \frac{1}{R_o} \qquad (3.31)$$

This implies

$$\ln \int_{\alpha}^{\beta} \frac{e^{-ra}\, p(a)m(a)da}{R_o} = -\ln R_o$$

Expanding e^{-ra} up to the term in r^2, we have

$$\ln \int_{\alpha}^{\beta} \left[1 - ra + \frac{r^2 a^2}{2}\right] \frac{p(a)m(a)}{R_o} \, da \simeq - \ln R_o$$

i.e.,

$$\ln \left[\int_{\alpha}^{\beta} \frac{p(a)m(a)da}{R_o} - r \int_{\alpha}^{\beta} a \frac{p(a)m(a)da}{R_o} + \frac{r^2}{2} \int_{\alpha}^{\beta} a^2 \, p(a)m(a)da\right.$$

$$= - \ln R_o$$

i.e.,

$$\ln \left[1 - r\mu + \frac{r^2}{2} (\mu^2 + \sigma^2) \right] \simeq -\ln R_o \tag{3.32}$$

Applying the expansion.

$$\ln (1 + y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \frac{y^4}{4} + - \ldots$$

equation (3.32) becomes

$$- r \mu + \frac{r^2}{2} (\mu^2 + \sigma^2) - \frac{(r \mu)^2}{2} \simeq - \ln R_o$$

which implies

$$\frac{\sigma^2}{2} r^2 - \mu r + \ln R_o \simeq 0$$

Therefore

$$r \simeq \frac{\mu - \sqrt{\mu^2 - 2\sigma^2 \ln R_o}}{\sigma^2} \tag{3.33}$$

Alternatively,

$$r \simeq \frac{\ln R_o}{\mu - \frac{\sigma^2}{2} r} \tag{3.34}$$

which can be solved iteratively by denoting the 'r' on the left hand side of (3.34) by r_{n+1} and the other by r_n. Thus we have

$$r_{n+1} = \frac{\ln R_o}{\mu - \frac{\sigma^2}{2} r_n} \tag{3.35}$$

for $n = 1, 2, 3, \ldots$

Coale's Method

Coale used the notion of the mean length of female generation T which is the number of years required to multiply the stable population by the average ratio of daughters to mother (NRR). That is

$$e^{rT} = NRR$$

Therefore

$$r = \frac{\ln NRR}{T} \tag{3.36}$$

The problem is now to determine T.

Let

$$\bar{m} = \frac{\int_{\alpha}^{\beta} am(a)da}{\int_{\alpha}^{\beta} m(a)da} \tag{3.37}$$

which is the mean age of child-bearing in a cohort subject to no mortality.

$$\mu_1 = \frac{\int_{\alpha}^{\beta} a\, p(a)m(a)da}{\int_{\alpha}^{\beta} p(a)m(a)da}$$

which is the mean age of child-bearing when mortality is taken into account.

$$\bar{a}_s = \frac{\int_{\alpha}^{\beta} a\, e^{-ra}\, p(a)m(a)da}{\int_{\alpha}^{\beta} e^{-ra}\, p(a)m(a)da} \tag{3.39}$$

is the mean age of child-bearing in a stable population. Expanding e^{-ra} in the numerator and denominator of formula (3.39), we have

$$\bar{a}_s = \frac{\int_{\alpha}^{\beta} a\left[1 - ra + \frac{r^2 a^2}{2} + \ldots\right] p(a)m(a)da}{\int_{\alpha}^{\beta} \left[1 - ra + \frac{r^2 a^2}{2} + \ldots\right] p(a)m(a)da}$$

$$= \frac{R_1 - R_2 r + R_3 \frac{r^2}{2} - R_4 \frac{r^3}{3!} + \ldots}{R_o - R_1 r + R_2 \frac{r^2}{2} - R_3 \frac{r^3}{3!} + \ldots}$$

$$= \frac{\frac{R_1}{R_o} - \frac{R_2}{R_o} r + \frac{R_3}{R_o} \frac{r^2}{2} - \frac{R_4}{R_o} \frac{r^3}{3!} + \ldots}{1 - \frac{R_1}{R_o} r + \frac{R_2}{R_o} \frac{r^2}{2} - \frac{R_3}{R_o} \frac{r^3}{3!} + \ldots}$$

By long division,

$$\bar{a}_s = \frac{R_1}{R_o} - \left[\frac{R_2}{R_o} - \left(\frac{R_1}{R_o}\right)^2\right] r + \left[\frac{R_3}{R_o} - \frac{3R_2 R_1}{R_o^2} + \frac{2R_1^3}{R_o^3}\right] \frac{r^2}{2} + \tag{3.40a}$$

$$= \mu_1 - (\mu_2 - \mu_1^2) r + (\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3) \frac{r^2}{2} + \ldots \tag{3.40b}$$

where

$$\mu_n = \frac{R_n}{R_o}$$

$$= \int_\alpha^\beta a^n \, f(a) da$$

$$= \int_\alpha^\beta a^n \frac{p(a)m(a)}{R_o} da$$

$$= E(a^n) \text{ , the nth moment.}$$

If we let

$$\lambda_1 = \mu_1 = E(a)$$

$$\lambda_2 = \mu_2 - \mu_1^2 = E(a - \mu_1)^2$$

$$\lambda_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = E(a - \mu_1)^3$$

etc.

then,

$$\bar{a}_s = \lambda_1 - \lambda_2 r + \lambda_3 \frac{r^2}{2} + \ldots \qquad (3.40c)$$

Alternatively, $\bar{a}_s$ can be expressed as

$$\bar{a}_s = \frac{\int_\alpha^\beta a e^{-ra} \, p(a)m(a)/R_o \, da}{\int_\alpha^\beta e^{-ra} \, p(a)m(a)/R_o \, da}$$

$$= - \frac{d}{dr} \ln \int_\alpha^\beta e^{-ra} \, p(a)m(a)/R_o \, da$$

$$= - \frac{d}{dr} \ln \int_\alpha^\beta e^{-ra} f(a) da \qquad (3.41)$$

where

$$f(a) = p(a)m(a)/R_o$$

which is a probability density function.

So $\int_\alpha^\beta e^{-ra} f(a)da$ is a moment generating function of r. Taking its log, we have a cumulant generating function.

Expanding e^{-ra} then

$$\bar{a}_s = - \frac{d}{dr} \ln \int_\alpha^\beta \left[1 - ra + r^2 \frac{a^2}{2} - r^3 \frac{a^3}{3!} + \ldots \right] f(a)da$$

$$= - \frac{d}{dr} \ln \left[1 - \mu_1 r + \mu_2 \frac{r^2}{2} - \mu_3 \frac{r^3}{3!} + \ldots \right]$$

Using the fact that

$$\ln (1+y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \frac{y^4}{4} + \ldots$$

then

$$\bar{a}_s = - \frac{d}{dr} \Big(- \mu_1 r + \mu_2 \frac{r^2}{2} - \mu_3 \frac{r^3}{3!} + \ldots \Big) - \tfrac{1}{2} \Big(- \mu_1 r + \mu_2 \frac{r^2}{2} - \mu_3 \ldots$$

$$+ \ldots \Big)^2 + \frac{1}{3} \Big(- \mu_1 r + \mu_2 \frac{r^2}{2} - \frac{\mu_1 + 3}{3} + \ldots \Big)^3 - \ldots$$

$$= \frac{d}{dr} \mu_1 r - (\mu_2 - \mu_1^2) \frac{r^2}{2} + (\frac{\mu_3}{6} - \frac{\mu_1 \mu_2}{2} + \frac{\mu_1^3}{3}) r^3 - \ldots$$

$$= \mu_1 - (\mu_2 - \mu_1^2) r + (\mu_3 - 3\mu_1\mu_2 + 2 \mu_1^3) \frac{r^2}{2} - \ldots$$

as in (3.40b)

$$= \lambda_1 - \lambda_2 r + \lambda_3 \frac{r^2}{2} - \ldots$$

as in (3.40c)

where λ_n (n = 1, 2, 3, ...) is the coefficient of $\frac{(-r)^2}{n^1}$ usually known as the nth cumulant of the net fertility schedule.

We also notice that

$$\bar{a}_s = - \frac{d}{dr} \ln \int_\alpha^\beta e^{-ra} \frac{p(a)m(a)da}{R_o}$$

$$= - \frac{d}{dr} \ln \frac{1}{R_o}$$

because of the characteristic equation

$$\int_\alpha^\beta e^{-ra} p(a)m(a)da = 1$$

Therefore

$$\bar{a}_s = \frac{d}{dr} \ln R_o$$

i.e.,

$$\ln R_o = \int \bar{a}_s dr$$

i.e.,

$$R_o = e^{\int \bar{a}_s dr}$$

i.e.,

$$NRR = e^{\int \bar{a}_s dr}$$

But NRR is also given by

$$NRR = e^{rT}$$

Therefore

$$r\,T = \int \bar{a}_s \, dr \tag{3.42}$$

i.e.,

$$r\,T = \int (\lambda_1 - \lambda_2 r + \lambda_3 \frac{r^2}{2} - \ldots) dr$$

$$= \lambda_1 r - \lambda_2 \frac{r^2}{2} + \lambda_3 \frac{r^3}{6} - \ldots$$

$$\simeq \lambda_1 r - \lambda_2 \frac{r^2}{2} \quad \text{(up to the term with } r^2\text{)}$$

Therefore

$$T \simeq \lambda_1 - \lambda_2 \frac{r}{2} \tag{3.43}$$

Therefore

$$r = \frac{\ln NRR}{T}$$

$$\simeq \frac{\ln NRR}{\lambda_1 - \lambda_2 \frac{r}{2}} \tag{3.44}$$

We should note that (3.44) is equivalent to (3.34) for λ_1 and λ_2 are the mean μ and various σ^2 respectively.

In the iterative fashion, we write

$$r_{n+1} = \frac{\ln NRR}{\lambda_1 - \frac{\lambda_2}{2} r_n}, \quad n = 1, 2, 3, \ldots$$

Coale used the first approximation as

$$r_1 = \frac{\ln NRR}{\lambda_1} = \frac{\ln NRR}{\mu_1} \tag{3.45}$$

So

$$r_2 = \frac{\ln NRR}{\lambda_1 - \frac{\lambda_2}{2\lambda_1} \ln NRR}$$

He let

$$\frac{\lambda_2}{2\lambda_1} = 0.8 \tag{3.46}$$

which he obtained from experimental data on various countries.

Therefore

$$r_2 = \frac{\ln NRR}{\lambda_1 - 0.8 \ln NRR} \qquad (3.47)$$

$$r_3 = \frac{\ln NRR}{\lambda_1 - \frac{\lambda_2}{2} r_2}$$

$$= \frac{\ln NRR}{\lambda_1 - \frac{\lambda_2}{2\lambda_1} \lambda_1 r_2}$$

$$= \frac{\ln NRR}{\lambda_1 (1 - 0.8 r_2)} \qquad (3.48)$$

and in general

$$r_{n+1} = \frac{\ln NRR}{\lambda_1 (1 - 0.8 r_n)} = \frac{\ln NRR}{\mu_1 (1 - 0.8 r_n)}$$

for n = 2, 3, 4, ...

Corollary:

The age distribution of a growing population is younger than that of the life table.

Proof:

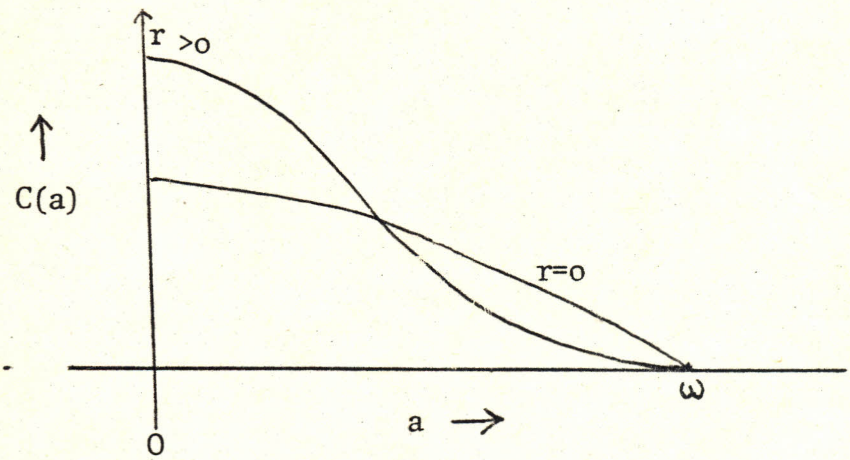
From (3.40a)

$$\bar{a}_s = \frac{R_1}{R_o} - \left[\frac{R_2}{R_o} - \left(\frac{R_1}{R_o}\right)^2\right] r + \ldots$$

$$= \frac{R_1}{R_o} - \sigma^2 r$$

$$< \frac{R_1}{R_o} \text{ when } r > 0. \qquad (3.49)$$

But $\frac{R_1}{R_o}$ is the mean age of a stationary population i.e. of a stable population with r equal to zero. The age distributions of a growing stable population and that of a stationary population is shown below:



3.7 How the intrinsic rate of increase affects birth rate, death rate and age distribution.

We wish to find how b, d and c(a) vary with r.

(i) $$b = \frac{1}{\int_0^w e^{-ra}\, p(a)da}$$

Therefore

$$\log b = -\log \int_0^w e^{-ra}\, p(a)da$$

$$\frac{d}{dr} \log b = -\frac{d}{dr} \log \int_0^w e^{-ra}\, p(a)da$$

i.e.,

$$\frac{1}{b} \cdot \frac{db}{dr} = \frac{1}{\int_0^w e^{-ra}\, p(a)da} \cdot \int_0^w ae^{-ra}\, p(a)da$$

$$= \bar{a}_s \quad \text{(as given in 3.39)}$$

Therefore

$$\frac{d\,b}{dr} = \bar{a}_s\, b \qquad (3.50)$$

$$\frac{d\,b}{dr} > 0$$

which implies that b is a non-decreasing function of r.

(ii) $c(a) = b\, e^{-ra}\, p(a)$

Therefore

$$\log_e c(a) = \log_e b - ar + \log_e p(a)$$

Therefore $\frac{d}{dr} \log c(a) = (\frac{d}{dr} \log b) - a$

$$= \frac{1}{b} \frac{d\,b}{dr} - a$$

$$= \bar{a}_s - a, \text{ cf } (3.50)$$

i.e.,

$$\frac{1}{c(a)} \frac{dc(a)}{dr} = (\bar{a}_s - a)$$

which implies

$$\frac{dc(a)}{dr} = (\bar{a}_s - a)\, c(a) \qquad (3.51)$$

So as r increases, c(a) increases for all ages less than $\bar{a}_s$, decreases for ages greater than $\bar{a}_s$ and does not change for $\bar{a}_s = a$.

(iii) $d = \int_0^w c(a)\, \mu(a)\, da$

Therefore

$$\frac{d}{dr} \ln d = \frac{d}{dr} \ln \int_0^w c(a)\mu(a)da$$

i.e.,

$$\frac{1}{d} \frac{dd}{dr} = \frac{1}{\int_0^w c(a)\mu(a)da} \cdot \int_0^w \frac{dc(a)}{dr} \mu(a)da$$

i.e.,

$$\frac{1}{d} \frac{dd}{dr} = \frac{1}{\int_0^w c(a)\mu(a)da} \int_0^w (\bar{a}_s - a)c(a)\mu(a)da$$

$$= \frac{1}{\int_0^W c(a)\mu(a)da}\left[\int_0^W \bar{a}_s\, c(a)\mu(a)da - \int_0^W a\, c(a)\mu(a)da\right]$$

$$= \bar{a}_s - \bar{a}_d$$

where

$$\bar{a}_d = \frac{\int_0^W a\, c(a)\mu(a)da}{\int_0^W c(a)\mu(a)da}$$

is the mean age of death.

Therefore

$$\frac{dd}{dr} = d(\bar{a}_s - \bar{a}_d) \qquad (3.52)$$

For max. or min.,

$$\frac{dd}{dr} = 0$$

which implies that $\bar{a}_d = \bar{a}_s$.

The second derivative is

$$\frac{d^2 d}{dr^2} = (\bar{a}_s - \bar{a}_d)\frac{dd}{dr} + d\left(\frac{d\bar{a}_s}{dr} - \frac{d}{dr}\bar{a}_d\right)$$

$$= d(\bar{a}_s - \bar{a}_d)^2 + d\left(\frac{d}{dr}\bar{a}_s - \frac{d}{dr}\bar{a}_d\right) \qquad (3.53)$$

But

$$\frac{d}{dr}\bar{a}_s = \frac{d}{dr}\frac{\int_0^W a\, e^{-ra}\, p(a)da}{\int_0^W e^{-ra}\, p(a)da}$$

$$= \frac{\left[\int_0^W e^{-ra}\, p(a)da\right]\left[-\int_0^W a^2\, e^{-ra}\, p(a)da\right] + \left[\int_0^W ae^{-ra}\, p(a)da\right.}{\left[\int_0^W e^{-ra}\, p(a)da\right]^2}$$

$$= - \frac{\int_0^w a^2 e^{-ra} p(a)da}{\int_0^w e^{-ra} p(a)da} + \bar{a}_s^2$$

$$= - \sigma_s^2 + \bar{a}_s^2$$

$$= \bar{a}_s^2 - \sigma_s^2 \qquad (3.54)$$

where σ_s^2 is the variance of the stable population

Similarly,

$$\frac{d}{dr} \bar{a}_s = \bar{a}_d^2 - \sigma_d^2 \qquad (3.55)$$

Therefore (3.5) becomes

$$\frac{d^2 d}{dr^2} = d(\bar{a}_s - \bar{a}_d)^2 + d\left[(\bar{a}_s^2 - \sigma_s^2) - (\bar{a}_d^2 - \sigma_d^2)\right]$$

$$= d(\bar{a}_s - \bar{a}_d)^2 + d\left[(\bar{a}_s^2 - \bar{a}_d^2) + (\sigma_d^2 - \sigma_s^2)\right] \qquad (3.56)$$

When $\bar{a}_s = \bar{a}_d$, then

$$\frac{d^2 d}{dr^2} = d(\sigma_d^2 - \sigma_s^2) \qquad (3.57)$$

So $\bar{a}_d$ is maximum if $\sigma_d^2 < \sigma_s^2$.

3.8 What happens to two stable Populations under given regimes of fertility and mortality schedules?

Theorem 3.1

Two stable populations with the same mortality schedule intersect at or about the mean of their mean ages.

Proof:

$$c(a) = b\, e^{-ra}\, p(a)$$

If the two stable populations have age distributions $c_1(a)$ and $c_2(a)$, then

$$\frac{c_1(a)}{c_2(a)} = \frac{b_1\, e^{-r_1 a}\, p(a)}{b_2\, e^{-r_2 a}\, p(a)}$$

$$= \frac{b_1}{b_2}\, e^{-(r_1 - r_2)a} \qquad (3.58)$$

At the intersection point,

$$\frac{c_1\,(a)}{c_2(a)} = 1$$

i.e.,

$$\frac{b_1}{b_2}\, e^{-(r_1 - r_2)a} = 1$$

which implies that

$$\hat{a} = \frac{\log b_1 - \log b_2}{r_1 - r_2} \qquad (3.59)$$

where â the horizontal axis of the intersecting point.
But we know from (3.50) that

$$\frac{d\,b}{dr} = \bar{a}_s b$$

which implies

$$\frac{1}{b}\,\frac{d\,b}{dr} = \bar{a}_s$$

which further implies that

$$\frac{d}{dr}\,\log b = \bar{a}_s$$

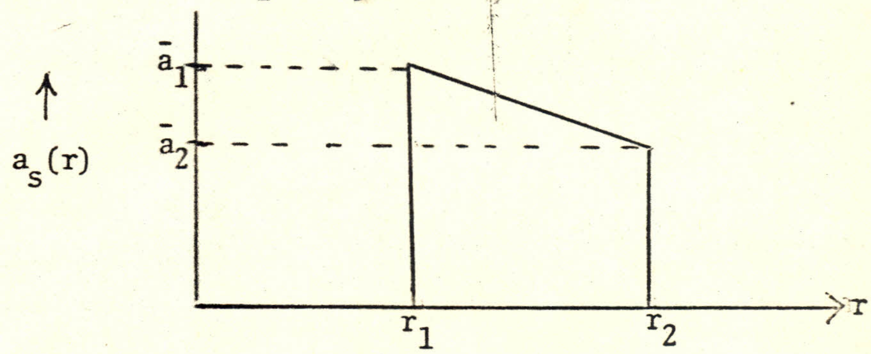
Therefore

$$\log b\Big|_{r_1}^{r_2} = \int_{r_1}^{r_2} \bar{a}_s \, dr$$

i.e.,

$$\log b_2 - \log b_1 = \int_{r_1}^{r_2} \bar{a}_s \, dr \tag{3.60}$$

If $\bar{a}_s$ is linear from r_1 to r_2 as shown in the diagram below,



then

$$\int_{r_1}^{r_2} \bar{a}_s \, dr = \text{the area of the trapezium}$$

$$= \frac{1}{2}(r_2 - r_1)(\bar{a}_1 + \bar{a}_2) \tag{3.61}$$

From (3.60) and (3.61), we therefore have

$$\frac{1}{2}(\bar{a}_1 + \bar{a}_2)(r_2 - r_1) = \log b_2 - \log b_1$$

which implies that

$$\frac{\bar{a}_1 + \bar{a}_2}{2} = \frac{\log b_2 - \log b_1}{r_2 - r_1} \tag{3.62}$$

Equating (3.59) and (3.62), we see that

$$\hat{a} = \frac{\bar{a}_1 + \bar{a}_2}{2}$$

Hence the proof.

Theorem 3.2

Two stable populations with a constant difference in mortality schedules at all ages combined with the same fertility schedule, have the same age distribution.

Proof:

We know that

$$c(a) = b\, e^{-ra}\, p(a)$$

$$= b\, e^{-ra}\, e^{-\int_0^a \mu(x)dx}$$

Let

$$c_1(a) = b_1\, e^{-r_1 a}\, p_1(a)$$

and

$$c_2(a) = b_2\, e^{-r_2 a}\, p_2(a)$$

Further let

$$\mu_2(x) = \mu_1(x) + k$$

where k is a constant.

Then

$$c_2(a) = b_2\, e^{-r_2 a}\, e^{-\int_0^a \mu_2(x)dx}$$

$$= b_2\, e^{-r_2 a}\, e^{-\int_0^a [\mu_1(a) + k]\, dx}$$

$$= b_2\, e^{-r_2 a}\, e^{-\int^a \mu_1(x)dx}\, e^{-ak}$$

$$= b_2\, e^{-r_2 a}\, p_1(a)\, e^{-ak}$$

Using the characteristic equation of the stable populations,

$$1 = \int_0^w e^{-r_1 a}\, p_1(a) m(a)\, da = \int_0^w e^{-r_2 a}\, p_2(a) m(a)\, da$$

This implies that

$$e^{-r_1 a}\, p_1(a) = e^{-r_2 a}\, p_2(a)$$

i.e.,

$$e^{+(r_2 - r_1)a} = \frac{p_2(a)}{p_1(a)}$$

$$= e^{-ak} \frac{p_1(a)}{p_1(a)} = e^{-ak}$$

This implies that

$$r_2 - r_1 = - k$$

i.e.,

$$r_2 = r_1 - k \qquad (3.63)$$

Therefore

$$c_2(a) = b_2 \, e^{-r_2 a} \, p_2(a)$$

$$= b_2 \, e^{-(r_1 - k)a} \, p_{2(a)}$$

$$= b_2 \, e^{-(r_1 - k)a} \, e^{-ak} \, p_1(a)$$

$$= b_2 \, e^{-r_1 a} \, p_1(a)$$

$$= b_1 \, e^{-r_1 a} \, p_1(a)$$

$$= c_1(a) \qquad (3.64)$$

So the two stable populations have the same fertility schedule,

$$b_2 = b_1 .$$

So we have shown a difference in mortality that would have no effect on the stable age distribution.

CHAPTER IV

A GENERALIZATION OF STABLE POPULATION RELATIONS

4.1 Relation between current population size and mortality schedule with age dependent growth rate.

Let $N(x)$ be the number of persons aged x, $\mu(x)$ the age specific mortality rate at exact age x, and r the constant growth rate.

In a stable population,

$$N(x) = N(o)\, e^{-rx}\, p(x) \qquad (4.1)$$

where $p(x)$ is the probability of surviving up to age x from birth.

Differentiating (4.1) with respect to x, we get

$$\frac{dN}{dx} = N(o)\left[- re^{-rx}\, p(x) + e^{-rx} \frac{dp}{dx} \right]$$

$$= N(o)\left[- re^{-rx}\, p(x) + e^{-rx} \frac{p(x)}{p(x)} \frac{dp}{dx} \right]$$

$$= N(o)\left[- re^{-rx}\, p(x) + e^{-rx}\, p(x) \frac{d}{dx} \log p(x) \right]$$

$$= N(o)\, e^{-rx}\, p(x) \left[- r + \frac{d}{dx} \log p(x) \right]$$

$$= N(x) \left[- r + \frac{d}{dx} \log p(x) \right] \qquad (4.2)$$

But

$$\mu(x) = - \frac{1}{\ell(x)} \frac{d\ell}{dx}$$

$$= - \frac{d}{dx} \log \ell(x)$$

$$= - \frac{d}{dx} \log \frac{\ell(x)}{\ell(o)} \cdot \ell(o)$$

$$= - \frac{d}{dx} \log \frac{\ell(x)}{\ell(o)}$$

$$= - \frac{d}{dx} \log \ p(x) \qquad (4.3)$$

Therefore (4.2) becomes

$$\frac{dN}{dx} = N(x) \left[- r - \mu(x) \right]$$

which implies that

$$\frac{1}{N(x)} \frac{dN}{dx} = - r - \mu(x) \qquad (4.4)$$

Thus the relative change in the number of persons at age x diminishes at a rate of $r + \mu(x)$.

Suppose now that the rate of increase is no longer a constant, but rather a function of age. Then equation (4.4) can be modified to

$$\frac{1}{N(x)} \frac{dN}{dx} = - r(x) - \mu(x)$$

i.e.,

$$\frac{d}{dx} \log N(x) = - r(x) - \mu(x) \qquad (4.5)$$

If $a \leq x \leq a+n$, then integrating (4.5), we have

$$\log N(x) \Big|_a^{a+n} = \int_a^{a+n} \left[- r(x) - \mu(x) \right] dx$$

which implies that

$$\log \frac{N(a+n)}{N(a)} = - \int_a^{a+n} r(x)dx - \int_a^{a+n} \mu(x)dx$$

i.e.,

$$N(a+n) = N(a) \, e^{- \int_a^{a+n} r(x)dx} \, {}_np_a \qquad (4.6)$$

where

$${}_np_a = e^{- \int_a^{a+n} \mu(x)dx} \qquad (4.7)$$

is the probability of surviving from age a to age a+n.

If $0 \leq x \leq a$, then we have

$$N(a) = N(o)\, e^{-\int_0^a r(x)dx}\, p(a) \qquad (4.8)$$

where

$$p(a) = e^{-\int_0^a \mu(x)dx}$$

is the probability of surviving from birth up to age a.

An alternative approach to obtaining the results derived above, is as follows:-

Let N(x,t) be the anumber of persons aged x at time t. Using the notion of total differentials,

$$dN(a,t) = \frac{\partial N(a,t)}{\partial a} da + \frac{\partial}{\partial t} N(a,t)dt \qquad (4.9)$$

At time t+dt, the number of persons aged x at time t who have died is

$$D(a,t) = N(a,t) - N(a+da, t+dt) \qquad (4.10a)$$

assuming closed population and the same cohort. Re-arranging (4.10a), we get

$$- D(a,t) = N(a+da, t+dt) - N(a,t) \qquad (4.10b)$$

By the principle of differential calculus, if

$$df = f(x+h, y+h) - f(x,y)$$

then

$$df = h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} \qquad (4.11)$$

as $(h, k) \to 0$.

So (4.10b) becomes

$$- D(a,t) = \frac{\partial N}{\partial a} da + \frac{\partial N}{\partial t} dt \qquad (4.12)$$

as $da = dt \to 0$

Therefore

$$- \frac{D(a,t)}{N(a,t)da} = \left(\frac{\partial N}{\partial a} + \frac{\partial N}{\partial t}\right) \frac{1}{N(a,t)}$$

i.e.,

$$- \mu(a,t) = \frac{1}{N(a,t)} \frac{\partial N}{\partial a} + \frac{1}{N(a,t)} \frac{\partial N}{\partial t}$$

But

$$\frac{1}{N(a,t)} \frac{\partial N}{\partial t} = r(a,t)$$

Therefore

$$- \mu(a,t) = \frac{1}{N(a,t)} \frac{\partial N}{\partial a} + r(a,t)$$

i.e.,

$$\frac{1}{N(a,t)} \frac{\partial N}{\partial a} = - r(a,t) - \mu(a,t)$$

i.e.,

$$\frac{\partial}{\partial a} \log N(a,t) = - r(a,t) = \mu(a,t)$$

as in (4.5). The results in (4.6) and (4.8) follow by integrating this equation if $a \leq x \leq a+n$ and if $o \leq x \leq a$ respectively.

4.2 Generalization of the equations characteristic of stable population.

The birth rate of the population is

$$b = \frac{N(o)}{\int_o^\infty N(a)da}$$

$$= \frac{N(o)}{\int_0^\infty N(o)\, e^{-\int_0^a r(x)da}\, p(a)da}$$

$$= \frac{1}{\int_0^\infty e^{-\int_0^a r(a)da}\, p(a)da} \qquad (4.13)$$

The proportion of the population that is age a is

$$c(a) = \frac{N(o)}{\int_0^\infty N(a)da}$$

$$= \frac{N(o)\, e^{-\int_0^a r(x)dx}\, p(a)}{\int_0^\infty N(o)\, e^{-\int_0^a r(a)da}\, p(a)da}$$

$$= \frac{e^{-\int^a r(a)dx'}\, p(a)}{\int_0^\infty e^{-\int_0^a r(x)dx}\, p(a)da}$$

$$= b\, e^{-\int_0^a r(x)dx}\, p(a) \qquad (4.14)$$

The birth rate can also be represented as

$$b = \int_a^\beta c(a)m(a)da$$

$$= \int_\alpha^\beta b e^{-\int_0^a r(x)dx}\, p(a)m(a)da$$

which implies that

$$1 = \int_\alpha^\beta e^{-\int_0^a r(x)dx}\, p(a)m(a)da \qquad (4.15)$$

where m(a) is the rate of bearing female children from women aged a.

4.3 Five - year age - groups.

For computational purposes, we shall consider 5-year age intervals along with some approximations. Proportion of people between ages x and x+5 is given by

$$_5C_x = \int_x^{x+5} C(a)\,da$$

$$= \int_x^{x+5} b\, e^{-\int_0^a r(y)dy}\, p(a)\,da \qquad (4.16)$$

using formula (4.14).

The mid-point of x and x+5 is x+2.5. So replace 'a' by x+2.5 in $e^{-\int_0^a r(y)dy}$. So we have

$$_5C_x \simeq \int_x^{x+5} be^{-\int_0^{x+2.5} r(y)dy}\, p(a)\,da$$

$$= b\, e^{-\int_0^{x+2.5} r(y)dy} \int_x^{x+5} p(a)\,da$$

$$= b\, e^{-\int_0^{x+2.5} r(y)dy}\, \frac{_5L_x}{\ell o} \qquad (4.17)$$

where

$$_5L_x = \int_x^{x+5} \ell(a)\,da$$

is the person-years lived between ages x and x+5

But

$$\int_0^{x+2.5} r(y)dy = \int_0^5 r(y)dy + \int_5^{10} r(y)dy + \int_{10}^{15} r(y)dy + \dots + \int_{x-5}^{x} r(y)dy + \int_x^{x+2.5} r(y)dy \qquad (4.18)$$

Assuming constant growth rates within the 5-year age interval so that $_5r_x$ becomes the growth rate between ages x and x+5, then (4.18) can be re-written as

$$\int_o^{x+2.5} r(y)dy = 5\left[{}_5r_o + {}_5r_1 + \ldots + {}_5r_2 + {}_5r_{x-5} \right] + 2.5({}_5r_x) \qquad (4.19)$$

Therefore

$$ {}_5C_x \simeq b\, e^{-\ 5({}_5r_o + {}_5r_1 + \ldots + {}_5r_{2.5}) + 2.5({}_5r_x)} \frac{{}_5L_x}{\ell o} \qquad (4.20)$$

From (4.17), we should note that

$$ {}_5C_o \simeq b\, e^{-\int_o^{2.5} r(y)dy} \frac{{}_5L_o}{\ell o}$$

$$ = b\, e^{-\ 2.5({}_5r_o)} \frac{{}_5L_o}{\ell o} \qquad (4.21)$$

To determine b, we sum (4.17) over x and get

$$\sum_{x=o}^{w} {}_5C_x \simeq b \sum_{x=o}^{w} e^{-\int_o^{x+2.5} r(y)dy} \frac{{}_5L_x}{\ell o}$$

But

$$\sum_{x=o}^{w} {}_5C_x = 1$$

Therefore

$$ b \simeq \frac{1}{\sum_{x=o}^{w} e^{-\int^{x+2.5} r(y)dy} \frac{{}_5L_x}{\ell o}} \qquad (4.22a)$$

$$ = \frac{1}{\sum_{x=o}^{w} e^{-\left[5({}_5r_o + {}_5r_5 + \ldots + {}_5r_{2.5}) + 2-5({}_5r_x)\right]} \frac{{}_5L_x}{\ell o}}$$

(4.22b)

Therefore

$$ {}_5C_o \approx \frac{e^{-2.5({}_5r_o)}\ {}_5L_o}{\sum_{x=o}^{w}\left[e^{-\int_o^{x+2.5} r(y)dy}\ {}_5L_x\right]} \tag{4.23a}$$

$$ = \frac{e^{-2.5\ ({}_5r_o)}\ {}_5L_o}{\sum_{x=o}^{w} e^{-\left[5({}_5r_o+{}_5r_5+\ ..+{}_5r_{x-5})+2.5({}_5r_x)\right]}\ {}_5L_x} \tag{4.23b}$$

and

$$ {}_5C_x \approx \frac{e^{-\int_o^{x+2.5} r(y)dy}\ {}_5L_x}{\sum_{x=o}^{w} e^{-\int_o^{x+2.5} r(y)dy}\ {}_5L_x} \tag{4.24a}$$

$$ = \frac{e^{-\left[5({}_5r_o+{}_5r_5\ +\ ...\ +\ {}_5r_{x-5})+2.5({}_5r_x)\right]}\ {}_5L_x}{\sum_{x=o}^{w} e^{-\left[5({}_5r_o+{}_5r_5+...+{}_5r_{x-5})+2.5({}_5r_x)\right]}\ {}_5L_x} \tag{4.24}$$

From two censuses, we can calculate the growth rate ${}_5r_a$ simply by applying the formula

$$ {}_5r_a = \frac{1}{t_2 - t_1}\ \ln \frac{({}_5N_a)t_2}{({}_5N_a)t_1} \tag{4.25}$$

where t_1 and t_2 are the periods censuses were taken. Note that formula (4.25) is from the fact that the population at time t growing at a rate r since x years ago

i.e.,

$$ N(t) = N(t - x)e^{rx} $$

which implies that

$$ r = \frac{1}{x}\ \ln \frac{N(t)}{N(t-x)} \tag{4.26}$$

For the open interval, say age A and over denoted by A^+, we have

$$C(A^+) \simeq b\, e^{-5({}_5r_o + {}_5r_5 + \ldots + {}_5r_{A-5}) - \frac{2}{3} e(A) r_{A^+}} \frac{T_A}{\ell o} \quad (4.27a)$$

$$= \frac{e^{-5({}_5r_o + {}_5r_5 + \ldots + {}_5r_{A-5}) - \frac{2}{3} e(A) r_{A^+}} T_A}{\sum_{x=o}^{w} e^{-\left[5({}_5r_o + {}_5r_5 + \ldots + 5\,{}_5r_{x-5}) + 2.5({}_5r_x)\right]} {}_5L_x} \quad (4.27b)$$

where

$e(A)$ = expectation of life at age A

r_{A^+} = the intrinsic growth rate of increase for ages A and over.

T_A = the total population for those aged A and over.

i.e.,

$T_A = \int_A^w \ell(a)\,da.$

We now wish to determine the ratio $\frac{T_x}{T_5}$ and $\frac{T_x}{T_{10}}$. These ratios are useful in detecting age mis-reporting.

From (4.23b) and (4.24b),

$$\frac{{}_5C_x}{{}_5C_o} = \frac{e^{-\left[5({}_5r_o + {}_5r_5 + \ldots + {}_5r_{x-5}) + 2.5({}_5r_x)\right]} {}_5L_x}{e^{-2.5({}_5r_o)} {}_5L_o}$$

$$= e^{-\left[2.5({}_5r_o) + 5({}_5r_5 + {}_5r_{10} + \ldots + {}_5r_{x-5}) + 2.5({}_5r_x)\right]} \frac{{}_5L_x}{{}_5L_o}$$

which implies that

$$\frac{{}_5L_x}{{}_5L_o} = \frac{{}_5C_x}{{}_5C_o} e^{2.5({}_5r_o) + 5({}_5r_5 + {}_5r_{10} + \ldots + {}_5r_{x-5}) + 2.5({}_5r_x)} \quad (4.28)$$

Once we have obtained a column of $\frac{{}_5L_x}{{}_5L_o}$ it is now a matter of adding these values from bottom upwards to get $\frac{T_x}{5L_o}$.

Hence

$$\frac{T_x}{T_5} = \frac{T_x}{{}_5L_o} \div \frac{T_5}{{}_5L_o} \qquad (4.29)$$

and

$$\frac{T_x}{T_{10}} = \frac{T_x}{{}_5L_o} \div \frac{T_{10}}{{}_5L_o}$$

From the office of Population Research, computerized tables of $\frac{T_x}{T_5}$ and $\frac{T_x}{T_{10}}$ against mortality levels have been made for each region (i.e. East, West, North and South) and sex. A graph of age against mortality level is then plotte.d It is hypothesized that if the graph is rising then there is an indication of over-statement of age. If the graph is showing a downward trend then we have under-estimation of age. Horizontal graph implies correct age statement.

4.4 Singulate Mean Age at Marriage (SMAM)

4.4.1 Derivation

The singulate mean at marriage (SMAM) is the mean age at first marriage among those who ever marry.

U(a) = the number of single persons at age a

N(a) = the total number of persons (all marital conditions) at age a.

Therefore, the proportion single at age a is

$$S(a) = \frac{U(a)}{N(a)} \qquad (4.30)$$

The proportion ever married at age a is

$$G(a) = 1 - S(a) \tag{4.31}$$

Let g(a) be the first marriage rate. Therefore

$$G(a) = \int_0^a g(x)dx \tag{4.32}$$

The first marriage distribution function can be constructed as

$$f(a) = \frac{g(a)}{\int_0^a g(x)dx}$$

i.e.,

$$f(a) = \frac{g(a)}{1-S(a)} \quad , \qquad 0 \leq a \leq A \tag{4.33}$$

$$= 0 \quad , \quad \text{otherwise}$$

So

$$E(a) = \int_0^A af(a)da$$

i.e.;

$$SMAM = \int_0^A af(a)da$$

$$= \frac{\int_0^A ag(a)da}{\int_0^A g(a)da}$$

$$= \frac{\int_0^A ag(a)da}{1 - S(A)} \tag{4.34}$$

where A is the greatest age of the first marriage.

Integrating by parts; i.e. letting

$$u = a \quad \text{and} \quad dv = g(a)da$$

implies

$$du = da \quad \text{and} \quad v = \int g(a)da = G(a)$$

Then

$$\int_0^A ag(a)da = uv\Big|_0^A - \int_0^A G(a)da$$

$$= aG(a)\Big|_0^A - \int_0^A G(a)da$$

$$= AG(A) - \int_0^A G(a)da$$

$$= A\left[1-S(A)\right] - \int_0^A \left[1-S(a)\right]da$$

$$= A-AS(A) - \left[A - \int_0^A S(a)da\right]$$

$$= A-AS(A) - A + \int_0^A S(a)da$$

$$= \int_0^A S(a)da - AS(A)$$

Therefore

$$SMAM = \frac{\int_0^A S(a)da - AS(A)}{1 - S(A)} \qquad (4.35)$$

Let d be the earliest age of marriage, then

$$SMAM = \frac{\int_0^a S(a)da + \int_d^A S(a)da - AS(A)}{1 - S(A)}$$

But

$$S(a) = 1 \quad \text{for } 0 \leq a \leq d$$

Therefore

$$\text{SMAM} = \frac{d + \int_{o}^{A} S(a)da - AS(A)}{1 - S(A)} \qquad (4.36)$$

In particular if

$$d = 15 \quad \text{and } A = 50$$

then

$$\text{SMAM} = \frac{15 + \int_{15}^{50} S(a)da - 50\ S(50)}{1 - S(50)} \qquad (4.37)$$

In the discrete form this would be

$$\text{SMAM} = \frac{15 + \sum_{15}^{45} 5^{S}x - 50\ S(50)}{1 - S(50} \qquad (4.38a)$$

$$= \frac{15 + 5 \sum_{i=1}^{7} S_i - 50\ S(50)}{1 - S(50)} \qquad (4.38b)$$

where 'i' is the ith age interval i.e. 15-19, 20-24, 25-29, ... , 45-49.

4.4.2 How to obtain SMAM using generalized stable population

Consider any aggregate of persons or objects with a continuous distribution by age since origin, or duration since event that defines membership. Then suppose a set of independent attrition factors (positive or negative: negative attrition could be people coming in by miracle, migration etc.) such that numbers leave the aggregate because of factor i at a rate $\mu_i(x)$.

Redistribution by age (duration) at a given moment has a given structure determined by arbitrary historical influences - wars, migration, varying births, deaths, miracles etc.

Under these conditons, then

$$N(a) = N(o)\, e^{-\int_o^a r(x)dx}\; e^{-\sum_i \int_o^a \mu_i(x)dx} \qquad (4.39)$$

where

$N(a)$ = Population size at age a

$r(a)$ = age specific growth rate

$\mu_i(x)$ = age specific attrition rate of factor i.

The number of single persons at age a is given by

$$U(a) = U(o)\, e^{-\left[\int_o^a j(x)dx + \int_o^a r_s(x)dx + \int_o^a \mu_s(x)dx\right]} \qquad (4.40)$$

where

$j(x)$ = the risk of first marriage

$r_s(x)$ = the age specific growth rate of single persons.

Therefore the proportion single is given by

$$S(a) = \frac{U(a)}{N(a)} = \frac{U(o)}{N(o)}\; \frac{e^{-\int_o^a j(x)dx - \int_o^a r_s(x)dx - \int_o^a \mu_s(x)dx}}{e^{\int_o^a r(x)dx - \int_o^a \mu(x)dx}} \qquad (4.41)$$

But

$U(o) = N(o)$

Therefore

$$S(a) = e^{-\int_o^a j(x)dx}\; e^{-\int_o^a \left[r_s(x) - r(x)\right]dx}\; e^{-\int_o^a \left[\mu(x)-\mu(x)\right]dx} \qquad (4.42a)$$

$$= S^*(a)\, e^{-\int_o^a \left[r_s(x) - r(x)\right]dx - \int_o^a \left[\mu_s(x)-\mu(x)\right]dx} \qquad (4.42b)$$

where

$$S^*(a) = e^{-\int_o^a j(x)dx}$$

which is the proportion single in a no-mortality cohort.

Thus

$$S^*(a) = S(a)\, e^{\int_0^a \left[r_s(x) - r(x)\right] dx + \int_0^a \left[\mu_s(x) - \mu(x)\right] dx}$$

and

$$SMAM = \frac{15 + \sum_{a=15}^{45} S(a) - 50\, S^*(50)}{1 - S^*(50)} \qquad (4.43)$$

In actual calculation, S(a) is considered as the average proportion single in two censuses

$$r_s(x) = \frac{1}{t_2 - t_1} \ \text{In}\ \frac{U(x,t_2)}{U(x,t_1)} \qquad (4.44)$$

Precisely

$$r_s(x) = \frac{U(x,t_2) - U(x,t_1)}{\text{Person-years lived}} \qquad (4.45)$$

or

$$r_s(x) = \frac{\text{Increase of single pop-S(I)-S(E)}}{\text{Mean population}} \qquad (4.46)$$

$$r_p(x) = \frac{\text{Increase of total pop-I-E}}{\text{Mean population}} \qquad (4.47)$$

4.5 <u>How to obtain migration rates using age specific growth rates</u>

We have already shown in the earlier section that

$$N(a) = N(o)\, e^{-\int_0^a r(x) dx}\, e^{\sum_i - \int_0^a \mu_i(x) dx}$$

where $\mu_i(x)$ is a decrement due to death, marriage, emmigration, etc.

Suppose

$$\mu_i(x) = -i(x)$$

which is emmigration at age x,

then

$${}_5N_a = {}_5N_o\, e^{-\int_o^a r(x)dx}\, e^{\int_o^a {}_5i_x dx}\, \frac{{}_5L_a}{{}_5L_o}$$

This implies

$$\int_o^5 {}_5i_x\, dx = \left[\ln\left(\frac{{}_5N_a}{{}_5N_o}\,\frac{{}_5L_o}{{}_5L_a}\right)\right] + \int_o^5 {}_5r_x\, dx \tag{4.48}$$

For

$$a = 5$$

we have

$$2.5({}_5i_o + {}_5i_5) = \left[\ln\left(\frac{{}_5N_5}{{}_5N_o}\cdot\frac{{}_5L_o}{{}_5L_5}\right)\right] + 2.5({}_5r_o + {}_5r_5) \tag{4.49}$$