SOME NEW DEMOGRAPHIC EQUATIONS IN SURVIVAL ANALYSIS UNDER GENERALIZED POPULATION MODEL: Applications to Swedish and Indian Census Age-data for Estimating Adult Mortality*

SUBRATA LAHIRI
Professor and Head
Department of Public Health & Mortality Studies
International Institute For Population Sciences,
Deonar, Mumbai-400 088
E-mail: sub_lahiri@vsnl.com

***Prepared for the 2004 Annual Meeting of Population Association of America (PAA) at Boston, Massachusetts, USA during April 1-3, 2004**

SOME NEW DEMOGRAPHIC EQUATIONS IN SURVIVAL ANALYSIS UNDER GENERALIZED POPULATION MODEL: Applications to Swedish and Indian Census Age-data for Estimating Adult Mortality

By

SUBRATA LAHIRI

[Abstract]:

This paper presents various formulas in estimating "10-year ***conventional*** and ***cumulative*** life table survival ratios", defined by the following ratios -- $_5L_{x+10}/_5L_x$ **and** T_{x+10}/T_x in life table terminology respectively, from two enumerations (not necessarily multiple of 5 years apart) of any closed population. The population under study should follow a generalized population model, and the age-specific growth curve should resemble closely to a second-degree polynomial. Attempts have also been made to establish algebraic relationships between census survival ratios (conventional and cumulative) and the corresponding life table survival ratios under GPM. The formulas, developed here, have been applied to a sufficiently accurate age-data of Sweden followed by those of India subject to serious response biases in age reporting. ***The proposed technique, which works quit well in assessing adult mortality even when the age-data are distorted due to age misreporting, may be extended for population projection and other demographic estimations.***

Key Words: ***Census survival ratios, Life table survival ratios, Destable population, and adult mortality***.

Introduction:

In census survivorship approach for life table construction from two enumerations (10 or 5 years apart) of a ***closed population*****,** it is conventionally assumed that the intercensal survivorship ratios are equal to the corresponding life table survival ratios that depict the average mortality experience during the intercensal period of the population under study. It has been shown in this paper that such a equality holds true ***only when*** the population under study is ***either stationary or stable***. A general relationship, applicable to any closed population (destabilized), between 10-year (*conventional*) census survival ratios (**10-CSRs**) and its associated 10-year (*conventional*) life table survival ratios (**10-LSRS**) has also been established under the assumption that the population under study follows a generalized population model of age-structure, and the intercensal period is ten years. A general formula has also been proposed for estimating **10-LSRs** from the age data of two consecutive censuses for any intercensal interval ***(not necessarily integral multiple of 5).***

In countries where census age-records are seriously affected due to age-misstatements, the conventional census survival ratios behave rather erratically and at times even exceed unity, which is absurd in a closed population. The use of cumulative census survival ratios, based on cumulated census age-returns beyond certain quinquennial ages as suggested by Coale and Demeny (United Nations, 1967), reduces such irregularities to a large extent. Thus, a general formula has also been developed in this paper in estimating '10-year ***cumulative*** life table survival ratios' **(10-cum-LSRs)**, defined by the ratio $\mathbf{T_{x+10}/T_x}$ in a life table terminology, from the corresponding '10-year ***cumulative*** census survival ratios' **(10-cum-CSRs)** under the assumption of generalized population model of age-structure applicable to any closed population having intercensal period of ***ten years***. A general equation has also been derived for estimating **10-cum-LSRs** from two consecutive census age-returns of a population having intercensal period ***other than 10 years not necessarily multiple of 5.***[1]

[1] In another study of the author (Lahiri, 2002) it has been shown how these estimated **10-cum-LSRs** as mentioned above can be used in locating an appropriate life table under a given model mortality pattern. It is worth noting that such a procedure **does not require** in projecting the initial population by age as enumerated in a particular census up to the next census date taken 5 or 10 years later than the former in contrast to the method proposed by Coale and Demeny (United Nations, 1967).

Methodology:

Notations used

In addition to the standard life table notations, the following notations will be used for the purpose of establishing the requisite relationship between population survival ratios **(PSRs)** and its associated life table survival ratios **(LSRs)**. The letters '**P**' and '**N**' are used to denote the number of persons related to a census enumeration and a mathematical model respectively. In subsequent discussions, the term 'age **x**' represents a discrete (integral) variable for a census but it stands for a continuous one in a mathematical model. The symbol **t** signifies the time variable--when **t = z** it stands for the first census date, and **t = z + 10** represents the second census date. Let,

$\mathbf{P_x(t)}$: Number of (enumerated) persons at age **x** **l.b.d** according to the census at time **t** **(t=z or z +10)**;

$\mathbf{{}_5P_x(t)}$: Number of persons in the age group (x, x+4) according to the census at time **t** **(t=z or z +10)**;

$\mathbf{P_{x+}(t)}$: Number of persons at ages '**x & above**' according to the census at time **t** **(t=z or z +10)**;

P(t) : Total number of persons according to the census at time **t (t=z or z +10)**.

$\mathbf{r, {}_5r_x \,\&\, r_{x+}}$: Average annual exponential growth rates during the decade (z, z+10) for the **whole population**, population in the **age-group (x, x+4)**, and sub population aged **x & above** respectively.

$\mathbf{N_x(t)}$: The number of persons **at exact age x at time t** according to the assumed mathematical population model of age-structure of the population under study. [Alternatively, $\mathbf{N_x(t)}$ may be treated as the frequency of the population model at time **t** such that the number of persons between the exact ages **x and x+Δx** at time **t**, **Δx** being the width of an infinitesimally small interval, is given by $\mathbf{N_x(t).\Delta x}$].

$\mathbf{{}_5N_x(t)}$: The number of persons between the exact ages x and x+5 at time t, according to the assumed population model and is given by:

$$\mathbf{{}_5N_x(t) = \int_x^{x+5} N_y(t)dy}$$

$N_{x+}(t)$: The number of persons aged x and above at time t in the assumed population model and is given by:

$$\mathbf{N_{x+}(t) = \int_x^w N_y(t)dy,}$$ **where '$\underline{w}$' is the maximum age that could be attained.**

$r_x(t)$: Instantaneous rate of growth of persons aged $\underline{\mathbf{x}}$ at time **t** according to the population model and is defined by,

$$\mathbf{r_x(t) = \frac{1}{N_x(t)} \frac{dN_x(t)}{dt}}$$

${}_{10}CS_x$: **10-year census survival ratio (10-CSR)** between the five-year age group (x, x+4) at time **t=z** and the age-group (x+10, x+14) at time **t=z+10**, and defined by the ratio **${}_5P_{x+10}(z+10)/{}_5P_x(z)$** which will be called, henceforth, in short, **10-CSR at age $\underline{x}$.**

${}_{10}PS_x$: **Ten-year population survival ratio (10-PSR)** between the five-year age-interval (x, x+5) at time **t=z** and the age-interval (x+10, x+15) at time **t=z+10**, and defined by the ratio **${}_5N_{x+10}$ (z+10)/${}_5N_x$(z)** which will be called, henceforth, **10-PSR at age $\underline{x}$**.

${}_{10}LS_x$: **Ten-year life table survival ratio (10-LSR)** between the five-year age-intervals (x, x+5) and (x+10, x+15) in the life table (or stationary) population associated with the population under study during the intercensal period (z, z+10), and defined by the ratio **${}_5L_{x+10}/{}_5L_x$** in life table terminology. In short, the above ratio will be called, henceforth, **10-LSR at age $\underline{x}$.**

${}_{10}CS'_{x+}$: **Ten-year cumulative census survival ratio (10-cum-CSR)** between the 'ages $\underline{x}$ & above' at time **t=z** and the 'ages x+10 & above' at time **t=z+10** and defined by the ratio **$P_{(x+10)+}$ (z+10)/P_{x+} (z)**, and in short the above ratio will be called, henceforth, **10-cum-CSR at age x.**

${}_{10}PS'_{x+}$: **Ten-year cumulative population survival ratio (10-cum-PSR)** between the 'ages $\underline{x}$ & above' at time **t = z** and the 'ages x+10 & above' at time **t = z+10**, and defined by the ratio **$N_{(x+10)+}$(z+10)/N_{x+}(z) which** will be called, henceforth, **10-cum-PSR at age $\underline{x}$**.

${}_{10}LS'_{x+}$: **Ten-year cumulative life table survival ratio (10-cum-LSR)** between 'ages x & above' and 'ages x+10 & above' in the life table (or stationary) population associated with the population under study during the intercensal period (z, z+10), and defined by the ratio **T_{x+10}/T_x** in life table terminology. In short, the above ratio will be called, henceforth, **10-cum-LSR at age** $\underline{x}$.

Relationship Between '10-PSR' and '10-LSR' at age 'a' in a Destabilized Population Model

According to a destabilized or generalized population model which is applicable to any population, the function $\mathbf{N_x(t)}$ describing the age-structure of any population at time **t** is given by the following equation (Bennett and Horiuchi, 1981; and see also, Preston and Coale, 1982):

$$\mathbf{N_x(t) = B(t)\exp\left[-\int_0^x r_y(t)\,dy\right]p(x;t)} \text{ ....}\mathbf{(1)}$$

where, **B(t)** : Number of births at time **t**; $\mathbf{r_y(t)}$: Instantaneous rate of growth of persons aged **y** at time t;

p(x;t) : Probability of surviving from birth to exact age **x** according to the stationary population associated with the destabilized population at time t and in life table terminology $\mathbf{p(x;t) = l_x(t)/l_0(t)}$, where $\mathbf{l_x(t)}$ denotes the number of survivors at exact age **x** out of the initial birth cohort $\mathbf{l_0}$ in the stationary population.

Given the age-distribution of persons of a destable population defined by the equation (1) at two points of time, ten-years apart, viz., the census dates **z** and **z+10**, our aim is to find a formula for estimating **10-LSR** at various quinquennial ages **x=a, a+5, a+10**, etc., from the corresponding values[2] of **10-PSR**. The set of estimates of **10-LSRs**, so obtained, depicts an average mortality experience of the population under study during the period (z, z+10) and thus may be treated as that related approximately to the time **z+5**, the mid-point of the period (z, z+10). Now, using the life table at time **z+5**, based on the set of **10-LSRs** values, which provides the value $\mathbf{p^*(x; z+5)}$, the average survival probability from birth to age **x** during the decade (z, z+10), one can find the analytical expression for the number of persons aged **x** at time z+5 in a destabilized population through the following equation:

$$\mathbf{N^*_y(z+5) \approx A^*(z+5) * \exp\left[-\int_0^x \bar{r}_y(z+5)\ dy\right] * l^*_x(z+5)} \text{-----}\mathbf{(2)}$$

where, $\mathbf{A^*(z+5) = B^*(z+5)\,/\,l^*_0(z+5)}$, and $\mathbf{B^*(z+5)}$ stands for the average annual number of births during the intercensal period (z, z+10) related approximately to the time **z+5**, the mid-point of the

---

[2] The observed value of '**10-PSR at x**' can be obtained through '**10-CSR at age x**' from the two census enumerations, ten years apart.

period (z, z+10). It can be shown that $\bar{r}_y(z+5)$ is nothing but the average annual exponential rate of growth of persons aged **y**, and $\mathbf{l_x^*}$ **(z+5)** represents the average survival function at age **x** during the period (z, z+10).

For the sake of simplicity in presenting various formulas, the 'argument' **z+5** within the parenthesis and the superscript -- asterisk sign **(*)** attached to various notations will be omitted, henceforth, in the subsequent development of different formulas. However, all the functions subsequently used here pertain approximately to the time **z+5** as an average experience during the decade (z, z+10) unless otherwise mentioned specifically.

Integrating both sides of (2) in the age interval (a, a+5) and using the first mean value theorem of Integral Calculus, we get,

$$\text{or } {}_5N_a \approx A * \exp\left[- \int_o^{a+2.5} \bar{r}_y \, dy \right] * {}_5L_a \qquad \text{.......... (3),}$$

$$\text{where } {}_5L_a = \int_a^{a+5} l_x \, dx \quad \text{by definition ....(3.1).}$$

Now, using (3), one may easily find that

$$\frac{{}_5N_{a+10}}{{}_5N_a} \approx \exp\left[- \int_{a+2.5}^{a+12.5} \bar{r}_y \, dy \right] * {}_{10}LS_a, \text{ for } a \geq 5 \text{ .... (4).}$$

The notation $_{10}$**LS**$_a$ **(defined by the ratio** $_5$**L**$_{a+10}$ / $_5$**L**$_a$ **in life table terminology)** used in the R.H.S of (4) stands for ten-year (***conventional***) life table survival ratio (**10-LSR) at age a** related to the time **z+5** that depicts the average survival experience during the decade (z, z+10). It may be noted that the ratio in L.H.S. of (4) is related to the time **z+5**. Thus the ratio $_5$**N**$_{a+10}$ /$_5$**N**$_a$ in the L.H.S. of (4) is not same as the **10-PSR at age a (**$_{10}$**PS**$_a$**)** during the decade (z, z+10), as the latter **(10-PSR)** is defined by the ratio $_5$**N**$_{a+10}$ **(z+10)/**$_5$**N**$_a$**(z)**. The problem of obtaining the relationship between **10-LSR at age a (**$_{10}$**LS**$_a$**)** and **10-PSR at age a (**$_{10}$**PS**$_a$**)** can be tackled in the following manner.

Denoting $_5$**r**$_x$ and $_5$**r**$_{x+10}$ as the average annual exponential rates of growth of persons aged (x, x+5) and (x+10, x+15) respectively during the time-interval (z, z+10), one may easily find the following approximations: $\mathbf{{}_5N_x(z+5) \approx \exp(5*{}_5r_x) * {}_5N_x(z)}$ **... (5)**

and $\mathbf{{}_5N_{x+10}(z+5) \approx \exp(-5*{}_5r_{x+10}) * {}_5N_{x+10}(z+10)}$ **....... (5.1).**

The quantities $\mathbf{{}_5r_x}$ **and** $\mathbf{{}_5r_{x+10}}$ in (5) and (5.1) can be estimated from the two consecutive census age-data, 10 years apart, through the following formula:

$$ {}_5\hat{r}_y = 0.1 * \ln\left\lfloor {}_5P_y(z+10)/{}_5P_y(z) \right\rfloor \text{, for } y = x \text{ \& } x+10 \text{ ....(5.2).}$$

Now, dividing (5.1) by (5) and simplifying the expression and **putting x = a**, we get,

$$\frac{{}_5N_{a+10}(z+5)}{{}_5N_a(z+5)} \approx {}_{10}PS_a * \exp\left[-5*\left({}_5r_a + {}_5r_{a+10}\right)\right] \quad \text{............. (6).}$$

The quantity $\mathbf{{}_{10}PS_a}$ **(defined by the ratio** $\mathbf{{}_5N_{a+10}(z+10)/{}_5N_a(z)}$ as mentioned earlier under the heading **'Notations used')** used in (6) stands for '**10 - PSR at a'** during the decade (z, z+10). Now, using (4) & (6), and re-arranging the terms, we get,

$$ {}_{10}LS_a = {}_{10}PS_a * \exp(R), \text{ for } \underline{a} > 0 \qquad \text{... (6.1).}$$

The symbol $\mathbf{{}_{10}LS_a}$, used in (6.1), stands for '**10 - LSR at age a'** and, **R** is given by,

$$R = -5*({}_5r_a + {}_5r_{a+10}) + \int_{a+2.5}^{a+12.5} \bar{r}_y \, dy \quad \text{.............(6.2).}$$

The equation (6.1) along with (6.2) provides the requisite relationship between **10-LSR at age a** and **10-PSR at age a**. The equation (6.1) clearly indicates that in general ***${}_{10}LS_a$ is not equal to ${}_{10}PS_a$*** excepting the case where **exp(R)** becomes unity which occurs ***only when R vanishes***. One may easily verify that the ***vanishing of R*** is ensured in the following situations ----**(i)** the population under study is stationary, that is $\bar{r}_a = 0$ ***for all a***, **(ii)** the population under study is perfectly stable, that is $\bar{r}_a = r$ ***(const.) for all a*** and **(iii)** $\bar{r}_a$ is a linear function of age **a** and of the form $\bar{r}_a = \mathbf{A + B*a}$, where A and B are non-zero constants.[3]

---

[3] It may be noted that the cases (i) and (ii) are the particular cases of (iii) where **A=B=0, and A≥0 but B=0** respectively. The linear growth curve under the case (iii) is **not** ordinarily expected in a destabilized population. Defining $\mathbf{r_x(t) = \frac{d}{dt}[\ln N_x(t)]}$, it can be shown that $\mathbf{N_x(t) = C(t)*\exp[a(t)+x*b(t)]}$, where **a(t), b(t) and c(t)** are non-zero constants under the case (iii) where $\mathbf{r_x = A(t)+B(t)*x}$. A close adherence to the form of $\mathbf{N_x}$ function mentioned above shows that such an age-structure is rather unusual in a destabilized population.

To use the formula (6.1) in practice for estimating ${}_{10}\mathbf{LS_a}$ from ${}_{10}\mathbf{PS_a}$ it is necessary to evaluate the integral in the R.H.S. of (6.2). Assuming $\bar{r}_x$ **- curve** being a second degree polynomial, the integral in the R.H.S. of (6.2) can be evaluated through Simpson's one-third rule of integration, and thus by using the above rule of integration, we get,

$$\int_{a+2.5}^{a+12.5} \bar{r}_x \, dx \approx \frac{5}{3}\left(\bar{r}_{a+2.5} + 4 * \bar{r}_{a+7.5} + \bar{r}_{a+12.5}\right) \text{ for } a \geq 5$$

$$\approx \frac{5}{3}\left({}_5r_a + 4 * {}_5r_{a+5} + {}_5r_{a+10}\right) \text{ for } a \geq 5 \text{ .... (7),}$$

where ${}_5r_x = \bar{r}_{x+2.5}$, for **x = a, a+5 & a+10**. Since the quantity $\bar{r}_{x+2.5}$ is the average annual exponential rate of growth of persons at (exact) age x+2.5, the mid-point of the age-interval (x, x+5), during the decade (z, z+10), the quantity $\bar{r}_{x+2.5}$ can be treated for all practical purposes approximately equal to ${}_5r_x$, the annual exponential rate of growth of persons aged (x, x+5).

Now, using **(6.2) & (7) in (6.1)** and simplifying the expression, we find

$${}_{10}LS_a \approx {}_{10}PS_a * \exp\left[-\frac{10}{3}\left({}_5r_a - 2 * {}_5r_{a+5} + {}_5r_{a+10}\right)\right], \ a \geq 5 \text{ ....(8)}$$

Using two consecutive decennial census age-returns presented in 5-year age-groups, one can easily obtain the estimate of **10-LSR at age <u>a</u>** (${}_{10}\mathbf{LS_a}$) by replacing ${}_{10}\mathbf{PS_a}$ by ${}_{10}\mathbf{CS_a}$, the intercensal survivorship ratio, and ${}_5r_a$ by ${}_5\hat{r}_a$, obtained through the formula (5.2) by replacing **y = a**. That is,

$${}_{10}L\hat{S}_a \approx {}_{10}CS_a * \exp\left[-\frac{10}{3}\left({}_5\hat{r}_a - 2 * {}_5\hat{r}_{a+5} + {}_5\hat{r}_{a+10}\right)\right], \text{ for } a \geq 5 \text{ ...(8.1)}$$

The formula (8.1) provides a procedure for translating 10-CSR at age <u>a</u> into the corresponding 10-LSR at age <u>a</u> in any closed population under the assumption that the age-specific growth curve $\bar{r}_x$ follows a second-degree polynomial in the age-interval (a, a+10). One may easily verify from the equation (8.1) that the equality between **10-CSRs** and **10-LSRs** holds good ***only when*** the population under study is either stationary (that is, $_5r_a = 0$ *for all a*) or stable (that is $_5r_a = r$ ***which is constant for all <u>a</u>***).

Estimation of 10-LSRs from Two Consecutive Census Enumerations with Intercensal Interval Other than 10-years and Not Necessarily Multiple of Five Years:

It may be noted that the formula (8.1) can be used when the intercensal period is exactly ten years. However, when the intercensal period is ***other than 10 years and not necessarily multiple of 5***, the following formula which is based on the equations (4) and (7) may be used for estimating **10-LSRs ($_{10}LS_a$)** :

$$_{10}LS_a \approx \frac{_5\overline{N}_{a+10}}{_5\overline{N}_a} * \exp\left[\frac{5}{3}\left(_5r_a + 4 * {_5r_a} + {_5r_{a+10}}\right)\right], \text{ for } a \geq 5 \text{ .... (9).}$$

The notation $_5\overline{N}_x$, **for x = a & a+10**, stands for average person-years lived by a person aged (x, x+5), and $_5r_x$**, for x = a, a+5 & a+10**, denotes the annual exponential rate of growth of persons aged (x, x+5) during the intercensal period (z, z+m) of **m** years, ***not necessarily integral multiple of five years***. The statistics $_5\overline{N}_x$ and $_5r_x$ can be estimated through the following formulas:

$$_5\hat{\overline{N}}_x = \frac{_5P_x(z+m) - {_5P_x(z)}}{m * {_5r_x}}, \text{ for } x = a \text{ and } a+10 \text{ ....(9.1)}$$

and

$$_5\hat{r}_x = \frac{1}{m}\ln\left[\left(_5P_x(z+m) / {_5P_x(z)}\right)\right], \text{ for } x = a, a+5 \,\&\, a+10 \text{ ....(9.2)}$$

The quantities $_5P_x(z)$, and $_5P_x(z+m)$ in (9.1) and (9.2) denote the number of person in the age-group (x, x+4) enumerated according to the censuses at time **z** and **z+m** respectively. The following explanations for using the equation (9.1) in estimating intercensal age-specific person-years lived, proposed by Preston and Bennett (1983), compared to the other existing procedures are worth noting.

The above approximation (equation 9.1) for $_5\overline{N}_x$, the average number of person-years lived[4] by a group of persons aged (x, x+5) during the intercensal period (z, z+m), makes use of the definition of an average annual rate of growth **($_5R_x$)** of persons aged (x, x+5) during the above mentioned period in terms of average annual increase in the number of persons aged (x, x+5) during

[4]Preston and Bennett (1983) proposed such an approximation for $_5\overline{N}_x$ while developing a technique for estimating adult mortality from two enumerations, m-years apart, under generalized population model of age-structure where the intercensal age-specific growth rates represent the average experience during the period.

the intercensal period (z, z+m) per person-years lived during the same period. Mathematically, ${}_5R_x$ can be defined by the following formula:

$$ {}_5R_x = \frac{{}_5P_x(z+m) - {}_5P_x(z)}{m * {}_5\overline{N}_x} $$

By inter-changing the variables ${}_5R_x$ **and** ${}_5\overline{N}_x$ in the above formula one could find an estimating formula (see the formula (9.3) as shown below) for ${}_5\overline{N}_x$ similar to (9.1) provided, of course, a reasonably good estimate of ${}_5R_x$ consistent with the two census enumerations is available.

$$ {}_5\overline{N}_x = \frac{{}_5P_x(z+m) - {}_5P_x(z)}{m * {}_5R_x} \quad \ldots\ldots\ldots\ldots\ldots(9.3) $$

Undoubtedly, the value of ${}_5\overline{N}_x$, estimated through the formula (9.3), depends upon whether the value of ${}_5R_x$ is equated to an arithmetic, or a geometric or an exponential rate of population growth. ***Since in a destable population in which number of persons in any age grows exponentially with the rate of growth applicable to that age, the quantity*** ${}_5R_x$ ***should be equated to*** ${}_5r_x$ ***, the average annual exponential rate of growth of persons aged (x, x+5) during the intercensal period, so as to obtain a reasonably good estimate of the average person-years during the intercensal period (z, z+m) – denoted by*** ${}_5\overline{N}_x$**.** One would like to know how far such an average person-years (${}_5\overline{N}_x$), estimated through the P-B method (equation 9.1), differs from those obtained by arithmetic mean -- A.M (${}_5\overline{N}_x^{(a)}$), and geometric mean -- G.M (${}_5\overline{N}_x^{(g)}$)**.** Keeping in mind that under a destable population in which the population aged (x, x+5) grows exponentially with annual growth rate ${}_5r_x$ during the intercensal period (z, z+m), that is ${}_5P_x(z+m) = {}_5P_x(z) * \exp(m * {}_5r_x)$, the values of ${}_5\overline{N}_x^{(a)}$ **and** ${}_5\overline{N}_x^{(g)}$**,** which are often used in demographic analysis, can be defined by the following equations,:

$$ {}_5\overline{N}_x^{(a)} = \frac{{}_5P_x(z)}{2} * \left[1 + \exp(m * {}_5r_x)\right] \quad \ldots\ldots\ldots(9.4) $$

$$ {}_5\overline{N}_x^{(g)} = {}_5P_x * \exp\left(\frac{m * {}_5r_x}{2}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots..(9.5) $$

Under the P-B method, the value of ${}_5\overline{N}_x$ is given by the following equation:

$$ {}_5\overline{N}_x = \frac{{}_5P_x(z)}{m * {}_5r_x} * \left(\exp(m * {}_5r_x) - 1\right) \quad \ldots\ldots\ldots\ldots(9.6) $$

Now, by expanding and rearranging the exponential function term by term in power of y_x ($=\mathbf{m{*}_5r_x}$) in ${}_5\overline{\mathbf{N}}_x$, ${}_5\overline{\mathbf{N}}_x^{(a)}$ **and** ${}_5\overline{\mathbf{N}}_x^{(g)}$ by Taylor's series it can be shown analytically that the estimate of person-years lived (${}_5\overline{\mathbf{N}}_x$) obtained through the formula (9.1) proposed by Preston and Bennett (P-B, 1983) lies between those of the estimates ${}_5\overline{\mathbf{N}}_x^{(g)}$ and ${}_5\overline{\mathbf{N}}_x^{(a)}$ obtained under geometric and arithmetic means respectively; and the value of ${}_5\overline{\mathbf{N}}_x$ can be approximated well through the following approximation: ${}_5\overline{\mathbf{N}}_x \approx \mathbf{a}{*}{}_5\overline{\mathbf{N}}_x^{(g)} + (1-\mathbf{a}){*}{}_5\overline{\mathbf{N}}_x^{(a)}$, **where** $\mathbf{a = 0.66667}$.

The above analytical exposition indicates that, in general, the P-B estimate for ${}_5\overline{\mathbf{N}}_x$ ***is not identical to that of the GM estimate (*** ${}_5\overline{\mathbf{N}}_x^{(g)}$ ***). However, they will be sufficiently close to each other in situations where*** ${}_5r_x$ ***is sufficiently small such that the term with higher powers of*** $y_x(=m.{}_5r_x)$ ***beyond the second degree in the Taylor's expansion of the exponential functions in the R.H.S of the equations (9.5) & (9.6) can be ignored for all practical purposes.***

There are some sophisticated and theoretically more precise techniques for estimating age-specific intercensal person-years lived which have been developed in the recent past (e.g., Coale, 1984; Bhat, 1987; and see also Bhat, 1995). However, in the context of poor quality of census returns in many developing countries whose age-sex-data are often distorted due to age-misreporting and digital preferences in age-reporting particularly at ages ending with **0** or **5**, we prefer to use the P-B method not only because of its operational convenience compared to the other two methods but also due to the fact that the P-B method is applicable ***even when the intercensal period is not an integral multiple of 5***. It may be noted here that the Coale's procedure requires sufficiently accurate single year age-returns,[5] and the technique proposed by Bhat (1987) though simpler than the Coale's approach is applicable ***only when*** the intercensal period is **10 or 5 years**[6]. Thus, in the context of poor quality of age-data, as frequently found in many developing countries, mere methodological sophistication under certain assumptions, which are unable to control (or at least dilute) the adverse

[5] Coale (1984) and Bhat (1987) proposed alternative methods for developing countries where single-year census age-returns are subjected to heaping at ages ending with digits zero, and/or five.

[6] Bhat's (1987) method assumes (i) that the population within a 5-year age interval is linearly distributed, and (ii) that deaths to a 5-year cohort are uniformly distributed during the intercensal period. Owing to the fact that the above assumptions, particularly the second one, do not hold good in general at younger and older ages (including open-ended terminal age-interval), the author proposed to make use of appropriate model life table to obtain suitable interpolation factors for estimating the person-years lived at those ages. The selected model mortality pattern, assumed to be applicable to the population under study, has considerable impact on the estimates.

effects of the error in the data, seems unwarranted.

Estimation of 10-LSRs & 5-LSRs under stability with intercensal period other than 10 years

When the population under study is stable or approximately stable with growth rate **r**, the formula (9) becomes: $${}_{10}LS_a \approx \frac{{}_5\overline{N}_{a+10}}{{}_5\overline{N}_a} * \exp(10 * r), \text{ for } a \geq 5 \quad \text{.....................} \quad (10)$$

Knowing the values of 10-LSR (${}_{10}LS_a$) through the equation (8) or (9) applicable to any closed population (destable) or through the equation (10) applicable to an approximately stable population, the values of 5-LSR (${}_5LS_a$) can be estimated through the following approximation:

$${}_5LS_a = [{}_{10}LS_{a-5} * {}_{10}LS_a]^{1/4}, \text{ for } 5<a<w-5 \quad \text{... (10.1)}$$

where **w**, being the initial age of the open-ended age interval, and

$${}_5LS_5 = {}_{10}LS_5 \,/\, {}_5LS_{10} \quad \text{... (10.2)}$$

and

$${}_5LS_{w-5} = {}_{10}LS_{w-10} \,/\, {}_5LS_{w-10} \quad \text{... (10.3)}$$

Relationship Between '10-cum-PSR at Age a' and the Associated '10-cum-LSR at age a' in a Destabilized Population Model

Integrating both sides of the equation (2) in the age-range (a, w) where **w** being the maximum age attainable by a person in the population under study, and using the first mean value theorem of integral calculus, there exists a point (age) C_{a+} lying between the ages **a** and **w** such that the following identify holds true:

$$N_{a+} \approx A * \exp\left[- \int_0^{C_{a+}} \bar{r}_y \, dy \right] * T_a, \text{ where } a< C_{a+} <w, \text{ and } T_a = \int_a^w l_x \, dx \quad \text{............} \quad (11)$$

Though the exact magnitude of C_{a+} is not known, however the mean age of the sub-population beyond age **a** may be taken as an approximation to C_{a+}. Some analytical justification of such an approximation can be found elsewhere (see, Lahiri, 1983, pp.143-148). It will be found later on that the magnitude of the difference $C_{(a+10)+} - C_{a+}$ is more important than those of the individual C_{a+}**'s** in determining relationship between **10-cum-PSR at age a** and **10-cum-LSR at age a** in any destabilized population. Now, using (11) we get,

$$\frac{N_{(a+10)+}}{N_{a+}} \approx \exp\left[-\int_{C_{a+}}^{C_{(a+10)+}} \bar{r}_y \, dy\right] * {}_{10}LS'_{a+} \text{ .... (11.1)}$$

where C_{a+} and $C_{(a+10)+}$ which are two points closely approximated by the respective mean-ages belonging to the age-intervals **'a & above'** and **'a+10 & above'** respectively and ${}_{10}LS'_{a+} = T_{a+10} / T_a$ **.** Since all the functions in (11.1) pertain to the time **z+5**, the ratio $N_{(a+10)+} / N_{a+}$ in the **L.H.S.** of (11.3) is also related to the time **z+5** and hence it is different from the **10-cum-PSR at age a** (${}_{10}PS'_{a+}$) during the period (z, z+10) which defined by is the ratio of persons **aged 'a+10 & above' at time z+10** to those **aged 'a & above' at time z**.

Following the similar procedure as adopted in developing the relationship between **'10-LSR at age a'** (${}_{10}LS_a$) and '**10-PSR at age a'** (${}_{10}PS_a$), shown under the formulas (4) to (6.1), the relationship shown below between **'10-cum-LSR at age a'** (${}_{10}LS'_{a+}$) and **'10-cum-PSR at age a'** (${}_{10}PS'_{a+}$) can be obtained by replacing ${}_5N_a(t)$ and ${}_5r_a$ by $N_{a+}(t)$ and r_{a+} respectively, and **R** by $I(C_{a+}, C_{(a+10)+})$:

$${}_{10}LS'_{a+} \approx {}_{10}PS'_{a+} * \exp\left[-5*\left(r_{a+} + r_{(a+10)+}\right)\right] * \exp\left[I\left(C_{a+}, C_{(a+10)+}\right)\right] \text{ ....(12),}$$

where, $I\left(C_{a+}, C_{(a+10)+}\right) = \int_{C_{a+}}^{C_{(a+10)+}} \bar{r}_y \, dy$ **............(12.1)**, and C_{a+} and $C_{(a+10)+}$ are two points closely approximated by the respective mean-ages belonging to the age-intervals **'a & above'** and **'a+10 & above'** respectively, and the notations ${}_{10}LS'_{a+}$ and ${}_{10}PS'_{a+}$ have already been defined under the section **"Notation Used".** To use the above relationship in practice it is necessary to evaluate the integral in the R.H.S. of (12.1). The integral denoted by $I(C_{a+}, C_{(a+10)+})$ can be evaluated numerically under certain assumptions regarding the nature of the $\bar{r}_y$ curve as discussed in the following section:

Evaluation of the Integral $I(C_{a+}, C_{(a+10)+})$ defined by (12.1)

Let $\mathbf{C_{a+}}$, $\mathbf{C_{(a+5)+}}$ **and** $\mathbf{C_{(a+10)+}}$ represent the mean ages of the sub-populations beyond ages '**a**', '**a+5**' and '**a+10**' respectively. If the whole interval of integration viz, $\mathbf{S \equiv (C_{a+}, C_{(a+10)+})}$ of width $\mathbf{h_{a+5}}$ is divided into two sub-intervals viz, $\mathbf{S_1 \equiv (C_{a+}, C_{(a+5)+})}$ and $\mathbf{S_2 \equiv (C_{(a+5)+}, C_{(a+10)+})}$ of width $\mathbf{h_{a+2.5}}$ and $\mathbf{h_{a+7.5}}$ respectively, and the two sub-intervals are almost of equal width (that is, $\mathbf{h_{a+2.5} \equiv h_{a+7.5}}$), then the integral **I** can be evaluated through the use of Simpson's one-third rule of numerical integration[7], under the assumption that the growth curve ($\bar{r}_y$) follows approximately a second degree polynomial in **y** within the whole range of integration denoted by **S**. Thus, applying the Simpson's one-third rule of numerical integration in (12.1), we get,

$$\mathbf{I\left(C_{a+}, C_{(a+10)+}\right) = \int_{C_{a+}}^{C_{(a+10)+}} \bar{r}_y dy = \frac{h^*_{a+5}}{3}\left[\bar{r}_{C_{a+}} + 4 * \bar{r}_{C_{(a+5)+}} + \bar{r}_{C_{(a+10)+}}\right]} \quad \text{.........(13)}$$

where $\mathbf{h^*_{a+5} = \left(h_{a+2.5} + h_{a+7.5}\right)/2}$ is the average width[8] of the two sub-intervals S_1 and S_2 having almost the same width. Similar to $\bar{r}_{x+2.5}$ which is well approximated by $\mathbf{{}_5r_x}$, the exponential rate of growth of persons aged (x, x+4) during the intercensal period, the value of $\bar{r}_{C_{x+}}$ may also be well-approximated by $\mathbf{r_{x+}}$, the exponential rate of growth of persons aged '**x & above**' during the intercensal period. An analytical explanation of the above logical argument that $\bar{r}_{C_{x+}}$ is being equal to $\mathbf{r_{x+}}$ can be found elsewhere (c.f. Lahiri, 1983; and Lahiri, 1985).

---

[7] The application of the Simpson's one-third rule of numerical integration is theoretically justified when the whole range of integration is sub-divided into **even** number of sub-intervals of **exactly equal width**. In case where the widths of the sub-intervals are **not** exactly equal but they are sufficiently close to each other, the rule may still be used to obtain an approximate value of the integral. However, the integrand should be approximately **second-degree polynomial** in the whole range of integration.

[8] The average width $\mathbf{h^*_{a+5}}$ of those of sub-intervals $\mathbf{S_1}$ **and** $\mathbf{S_2}$ is identical to the half of $\mathbf{h_{a+5}}$, the width of the whole interval $\mathbf{S\ (=S_1+S_2)}$.

Thus, by replacing $\bar{r}_{C_{x+}}$ **by r_{x+}, for x=a, a+5 & a+10 in (13)**, we get the value of **$I[C_{a+}, C_{(a+10)+}]$** as shown below:

$$I\left(C_{a+}, C_{(a+10)+}\right) \approx \frac{h^*_{a+5}}{3} * \left[r_{a+} + 4 * r_{(a+5)+} + r_{(a+10)+}\right] \text{ ....(13.1)}$$

It is worthwhile to mention here that the formula (13.1) produces sufficiently reliable estimate of the integral **I** provided the widths of the two sub-intervals **S_1** and **S_2** are sufficiently close to each other. However, when the widths of the two sub-intervals differ widely from each other, the following procedure may be adopted. If **k** be the mid-point of the interval **$(C_{a+}, C_{(a+10)+})$**, then the integral **I** can be obtained through the Simpson's rule as follows:

$$I\left(C_{a+}, C_{(a+10)+}\right) \approx \frac{h^*_{a+5}}{3} * \left[r_{a+} + 4 * \bar{r}_k + r_{(a+10)+}\right] \text{ .... (13.2)}$$

The value of $\bar{r}_k$ may be obtained as follows:

(i) If **k** belongs to the sub-interval **S_1**, that is **$C_{a+} < k < C_{(a+5)+}$** , then

$$\bar{r}_k = \frac{1}{h_{a+2.5}}\left[\left(C_{(a+5)+} - k\right) * r_{a+} + \left(k - C_{a+}\right) * r_{(a+5)+}\right] \text{ .... (13.3)}$$

where **$h_{a+2.5} = C_{(a+5)+} - C_{a+}$** is the width of the sub-interval **S_1**; or **(ii)** if **k** belongs to the sub-interval **S_2**, that is, **$C_{(a+5)+} < k < C_{(a+10)+}$**, then

$$\bar{r}_k = \frac{1}{h_{a+7.5}}\left[\left(C_{(a+10)+} - k\right) * r_{(a+5)+} + \left(k - C_{(a+5)+}\right) * r_{(a+10)+}\right] \text{ .....(13.4)}$$

where **$h_{a+7.5} = C_{(a+10)+} - C_{(a+5)+}$** represents the width of the sub-interval **S_2**.

A general relationship between ${}_{10}LS'_{a+}$ and ${}_{10}PS'_{a+}$ can be obtained by using (13.2) in (12), and re-arranging the terms we get,

$${}_{10}LS'_{a+} \approx {}_{10}PS'_{a+} * \exp\left[-\frac{2}{3}\left\{\left(15 - h^*_{a+5}\right) * r^*_{(a+5)+} - 2 * h^*_{a+5} * \bar{r}_k\right\}\right] \text{....... (14)},$$

where, $h^*_{a+5} = \frac{1}{2}\left[C_{(a+10)+} - C_{a+}\right]$, and $r^*_{(a+5)+} = \frac{1}{2}\left(r_{a+} + r_{(a+10)+}\right)$. It may be noted that h^*_{a+5} can also be regarded as the average widths of the two sub-intervals **S_1** and **S_2**.

The quantity $\bar{r}_k$ is given by (13.3) or (13.4) depending upon whether **k** belongs to **S_1 or S_2**. One may easily verify from (13.3) or (13.4), if **S_1 and S_2** are exactly of equal width, that is, when **k** coincides with $C_{(a+5)+}$, then $\bar{r}_k$ will be exactly equal to $r_{(a+5)+}$. Hence, when the widths of **S_1 and S_2** are of equal size, $\bar{r}_k$ in (14) should be replaced by $r_{(a+5)+}$ and h^*_{a+5} should be replaced by $\frac{1}{2} h_{a+5}$, half of the width of the interval **S**.

It is worthwhile to mention here that in the case of usual (or conventional) survival ratios, the equality between 10-LSRs and 10-PSRs holds good when the population under study is either stationary or stable; whereas in the case of cumulative survival ratios, the equality between 10-cum-LSRs and 10-cum-PSRs holds good only when the population under study is a stationary one. When the population under study is ***stable or approximately stable***, the equation (14) becomes:

$$_{10}LS'_{a+} = {}_{10}PS'_{a+} * \exp\left[-\left(10 - h_{a+5}\right) * r\right] \qquad \textbf{... (14.1)}$$

where **r** is the rate of growth of the stable or approximately stable population, and $h_{a+5}\ (= 2 * h^*_{a+5})$ $= C_{(a+10)+} - C_{(a+5)+}$. For obtaining the estimate of $_{10}LS'_{a+}$ from two enumeration, ten years apart, $_{10}PS'_{a+}$ in (14) or (14.1) should be replaced by its observed value $_{10}CS'_{a+}$, obtained from the age-data of the two consecutive decennial censuses.

<u>Estimation of 10-cum-LSRs from two consecutive census enumerations with intercensal interval other than 10-years</u>:

The formulas derived above for estimating **10-cum-LSRs** from the corresponding **10-cum-PSRs** are applicable when the intercensal interval is **10-years**. However, when the intercensal interval is other than 10 years **not necessarily multiple of 5**, the following formula obtained through a similar approach to that of **10-LSRs** shown under the equation (9), may be used for estimating **10-cum-LSRs**. The formula given below is based on the equations (11.1) and (13):

$$_{10}LS'_{a+} = \frac{\overline{N}_{(a+10)+}}{\overline{N}_{a+}} * \exp\left[\frac{h^*_{a+5}}{3} * \left(r_{a+} + 4 * \bar{r}_k + r_{(a+10+)}\right)\right], \textbf{ for } a \geq 5 \textbf{ .... (15).}$$

The value of $\overline{N}_{x+}$, for x = a & a+10 in R.H.S of (15) can be estimated as follows:

$$\overline{N}_{x+} = \sum_{y=x}^{w-5} {}_5\overline{N}_y + \overline{N}_{w+} \text{ for } x = a \ \& \ a+10 \ \ldots\ldots\ldots(15.1),$$

where, $${}_5\overline{N}_y \approx \frac{{}_5P_y(z+m) - {}_5P_y(z)}{m * {}_5\hat{r}_y} \ldots\ldots\ldots\ldots(15.2)$$

and

$$\overline{N}_{w+} \approx \frac{P_{w+}(z+m) - P_{w+}(z)}{m * \hat{r}_{w+}} \ldots\ldots\ldots\ldots\ldots(15.3)$$

The symbol **'w'** stands for initial age of the terminal open-ended age-interval, and estimates of $\mathbf{r_{x+}}$ **for** **x = a, a+10, & w** are given by: $\hat{r}_{x+} = \frac{1}{m}\left[\ln P_{x+}(z+m)/P_{x+}(z)\right] \ldots\ldots\ldots\ldots(15.4)$

An estimate of $\bar{r}_k$ can be obtained through the formula (13.3) or (13.4), and $\mathbf{h^{*}_{a+5}}$ is given by

$$h^{*}_{a+5} = \frac{1}{2}\left[C_{(a+10)+} - C_{a+}\right] \ ,$$

In this case of stable population with growth rate **r**, the equation (15) becomes

$${}_{10}LS'_{a+} = \frac{\overline{N}_{(a+10)+}}{\overline{N}_{a+}} \exp(h_{a+5} * r) \ \ldots\ldots\ldots(15.5),$$

where, $\mathbf{h_{a+5}}$ is the width of the interval $\mathbf{[C_{a+}, C_{(a+10)+}]}$.

Estimation of mean age beyond age a (C_{a+}) and h^{*}_{a+5}

While obtaining ${}_{10}LS'_{a+}$ through the equation (14) or (15), we need to know the value of $\mathbf{h^{*}_{a+5}}$, the ***half of the difference*** between mean ages of persons ***beyond ages*** **a+10** and **a**. The mean-age of persons aged **'a & above'** is generally obtained as the weighted average of the mean ages of all the 5-year age-intervals, that is their mid-points[9] together with an arbitrarily fixed mean-age of the open-ended terminal age-interval, taking the respective population sizes in various age-intervals including the terminal open-ended age-interval as the weights. ***In demographic estimation particularly for countries with limited and defective age data, the element of arbitrariness in locating the mean-age of an open-ended terminal age-interval is a vexing issue. To remove the above element of arbitrariness as much***

[9] Under the assumption of uniform distribution of persons over the 5-year age-interval (x, x+5), where x ≥ 5, the mean-age of persons aged (x, x+5) can be assigned to its mid-point, that is, x+2.5.

as possible we follow the procedure proposed by Preston and Lahiri (1991). For the benefit of the readers, the necessary technical aspects relating to the estimation of $\mathbf{C_{a+}}$, mean age of persons aged '**a & above'**, have been discussed in the **Technical Appendix A**.

Splitting of 10-cum-LSRs ($_{10}LS'_{a+}$) into 5-cum-LSRs ($_{5}LS'_{a+}$)

After estimating **10-cum-LSRs ($_{10}LS'_{a+}$)** through the formula (14) or (14.1), one can estimate **5-cum-LSRs ($_{5}LS'_{a+}$'s)** through the following approximations:

$$_{5}LS'_{(a+5)+} \approx (\,_{10}LS'_{a+} * _{10}LS'_{(a+5)+})^{\frac{1}{4}}, \text{ for } a = 5, 10, \ldots, w - 20 \ldots\ldots\ldots\ldots\ldots(15.6)$$

$$_{5}LS'_{5+} = \frac{_{10}LS'_{5+}}{_{5}LS'_{10+}} \ldots\ldots(15.7) \qquad \text{and} \qquad _{5}LS'_{(w-10)+} = \frac{_{10}LS'_{(w-15)+}}{_{5}LS'_{(w-15)+}} \ldots\ldots\ldots(15.8)$$

The symbol '**w**' stands for the initial age of the open ended terminal age-interval. The formula (15.6) is ***not only operationally convenient***, apart from its theoretical simplicity, in splitting **10-cum-LSRs** ($_{10}LS'_{a+}$) into **5-cum-LSRs ($_{5}LS'_{a+}$)** ***but also produces reasonably good estimates*** of **5-cum-LSRs** even when the values of **10-cum-LSRs** are estimated from the distorted census age-returns. This is because of the self-smoothening property of the formula (15.6) (for further details see Lahiri, 1983; pp. 310-324).

Two Mathematical Identities among Life Table Functions in Obtaining 5-cum-LSRs from 10-cum-LSRs (along with 10-LSRs), and 5-LSRs from 5-cum-LSRs alone

Knowing the values of **10-LSRs ($_{10}LS_a$) and 10-cum-LSRs ($_{10}LS'_{a+}$)** from the two enumerated age-returns of a closed destable population, not necessarily 10 years (or integral multiple of 5-years) apart, through the formulas (8.1) [or (9)] and (14) [or (15)], derived in this papers, one can obtained the values of **5-cum-LSRs ($_{5}LS'_{a+}$), and 5-LSRs ($_{5}LS_a$)** through the following ***identities*** respectively:

$$_{5}LS'_{a+} = \frac{_{10}LS_a - _{10}LS'_{a+}}{_{10}LS_a - _{10}LS'_{(a+5)+}} \ldots\ldots\ldots(16) \qquad \text{and} \qquad _{5}LS_a = \frac{_{5}LS'_{a+} * \left[1 - _{5}LS'_{(a+5)+}\right]}{1 - _{5}LS'_{a+}} \ldots\ldots\ldots\ldots(16.1)$$

One can easily verify the above two identities by replacing the survival ratios $_{10}LS_a$, $_{5}LS'_{a+}$ **and** $_{10}LS'_{a+}$ with their corresponding ***standard*** life table functions as mentioned under the *sub-heading* ***'notations used'***

In the presence of heavy age misreporting the values of **5-cum-LSRs in contrary to 5-**

LSRs, estimated through the procedure mentioned in the earlier paragraphs, lie between 0 and 1 for all ages, and its **(5-cum-LSRs)** visual examination generally show a smooth and regular declining pattern. The errors in age-reporting become almost invisible (or latent) due to dampening effect of cumulation. However, the values of **5-LSRs ($_5LS_a$),** which can be obtained through the formulas (16.1) after estimating **5-cum-LSRs ($_5LS'_{a+}$)** through the procedure mentioned earlier in this paper, are likely to show some erratic pattern[10]. Further investigations may be carried out in obtaining adjusted series of **5-LSRs** by smoothing the ***latent irregularities*** in **5-cum-LSRs**, as mentioned above, through a suitable mathematical graduation formula, such as -- Brass type two-parameter logit model . It is worth mentioning in this context that the smoothing procedures for census age-data or the raw (***conventional***) census survival ratios, distorted due to age-misreporting, are rather arbitrary and often influenced by personal predilections. Thus, such an attempt of graduating the estimated **5-cum-LSRs** through suitable mathematical model would be particularly helpful in obtaining smooth series of **5-LSRs** without smoothening the distorted age-data or the raw survival ratios (see, Lahiri, 1983 for further discussions). It has been shown elsewhere that knowing the values of **5-cum-LSRs ($_5LS'_{a+}$'s)** beyond age 5, defined by the ratio of the form T_{x+5} / T_x in life table terminology, how one can estimate an appropriate set of 5-year survival probabilities ($_5p_x$) beyond age 5 consistent with the **5-cum-LSRs** values under certain assumptions regarding the nature of the **l_x-curve** (for details see, Lahiri, 1985 & 2003).

Some Methodological Issues of the Present Method Compared to the UN Classical Forward and Backward Projection and Cumulation Method and Preston-Bennett Method for Estimating Intercensal Mortality

In the context of heavy response biases in age-reporting due to digital preference and/or ignorance of correct age in many developing countries, the census based ***observed*** values of **10-CSRs ($_{10}CS_a$'s)** or the values of **10-LSRs ($_{10}LS_a$),** estimated through equation (8.1), behave very erratically, and at times they exceed unity, which is rather absurd in a closed population. To avoid this difficulty,

---

[10] This is ***not only*** because of age-misreporting in the age-returns but also due to the assumptions involved in translating continuous form of the formulas into the corresponding discrete forms in estimating **10-LSRs and 10-cum-LSRs** from two age-returns of a closed population undergoing generalized population model.

one may make use of cumulated populations beyond ages 5, 10, 15, etc., so as to calculate the 10-year cumulative census survival ratios **(10-cum-CSRs)** as suggested by Coale and Demeny (United Nations, 1967; see also United Nations, 1983)). ***Since no formula was available earlier to translate directly '10-year cumulative census survival ratios' into the corresponding '10-year cumulative life table survival ratios' applicable to a destable population,*** Coale and Demeny proposed to project the first census age-distribution forward repeatedly using a family of model life tables at various levels up to the time of the second census taken ten years latter. The authors finally suggested that the appropriate model mortality table could be identified by comparing the projected population beyond a given age to that of the enumerated population in the second census. While assessing intercensal adult mortality in a destable population through the application of the classical forward and backward projection methods, some researchers (Palloni Kominsky, 1984; Bhat, 1995) found a significant difference in the results and they pointed out that the presence of reporting errors in the enumerated age-data is primarily responsible for different results between the two procedures. It is worthwhile to mention here that ***such a disagreement between the two procedures may occur even when the age-data are sufficiently accurate. This is primarily because of the fact that the method of classical forward projection and that of backward projection in estimating intercensal adult mortality inherently assume the equality between conventional census survival ratios, and the corresponding conventional life table survival ratios. It has been shown earlier that such equality holds good only when the population under study is either stationary or stable. However, in the case of cumulative survival ratios one may easily verify that such equality holds good only when the population under study is stationary as shown by the formula (14).***

In nutshell the present technique, in contrast to the Coale-Demeny method, neither requires the assumption of equality between **10-CSRs** and **10-LSRs** nor requires projecting the initial age-distribution of persons under different mortality levels up to the end of 10 years period. The method proposed here is applicable even when the intercensal period is other than 10 years (not necessarily multiple of 5). In the presence of heavy response biases in age reporting as frequently found in many developing countries, the use of **10-cum-CSRs** which in turn produce **10-cum-LSRs** under generalized population model reduces the errors due to age-misreporting to a large extent. Furthermore, the direct use of cumulated age-data in the formulas developed here for estimating intercensal mortality has certain advantages over that of Preston-Bennett (1983) in controlling the error due to age-misreporting.

In discrete terms, the life expectancy at age x (**e^0_x**) based on the basic equations for constructing census based adult mortality table, proposed by Preston and Bennett (1983) can be expressed by the following approximate formula (see also, United Nations, 1983):

$$e^0_x \approx \frac{\sum_{y=x}^{w} {}_5\overline{N}_y * \exp\left[5.0 \sum_{u=x}^{y-5} {}_5 r_u + 2.5 * {}_5 r_y\right]}{\overline{N}(x)} \text{ ---------------- (17)}$$

where $\overline{N}(x)$, the number of persons at exact age x, was be estimated through the following approximation as proposed Preston and Bennett (1983):

$$\hat{\overline{N}}(x) \approx \frac{{}_5\hat{\overline{N}}_{x-5} * \exp[-2.5 * {}_5\hat{r}_{x-5}] + {}_5\hat{\overline{N}}_x * \exp[2.5 * {}_5\hat{r}_x]}{10} \text{ ----------(17.1)}$$

where the notations used in (17) and (17.1) have already been explained in the text.

Comparing the formulas (14) and (15) to that of (17), mentioned above, one may easily find that the values of **10-cum-LSRs**, estimated through the formula (14) or (15), make direct use of cumulated population whereas the numerator of (17) makes use of weighted sum of the form **$\sum {}_5N_y.\exp[R(y)]$**, (where **R(y)** stands for the quantity under the exponential sign in the equation (17), instead of $\sum {}_5N_y$ as used in the formula (14) or (15). Furthermore, the development of the formula (14) or (15) is based on the assumption of second degree polynomial of the growth curve ($\overline{r}_x$) in contrast to the formula (17) that assumes the linearity of $\overline{r}_x$-curve within each of the 5-year age-intervals ***excluding*** the terminal open-ended age-interval.

Application to the Age-data of Sweden and India for Estimating Longevity at Adult Ages

I. Swedish Females, 1966-70

Swedish data are well known for their accuracy and have often been used by demographers to examine new estimation procedure. In the present investigation, the proposed technique was applied to the age-data of Swedish females during 1966-70. The mean person-years lived during 1966-70 by quinquennial ages together with the respective age-specific growth rates were borrowed from a study carried by Preston and Bennett (1983). The basic data and the major steps for estimating $_{10}\mathbf{LS}'_{\mathbf{a}}$ values from census age-returns under the generalized population model are shown in Table-1.The overall mortality level ($\mathbf{e^0_0}$) consistent with the estimated $_5\mathbf{LS}'_{\mathbf{a+}}$ values (shown in Table 2) under the Coale and Demeny (C-D) West model mortality pattern and its associated life expectancies at various ages are also shown in Table 2.

As the accuracy of the Swedish census age-reporting is well-recognized, one would expect that the mortality level ($\mathbf{e^0_0}$) associated with each of the $_5\mathbf{LS}'_{\mathbf{a+}}$ values should be almost identical to each other. But the estimates of mortality level ($\mathbf{e^0_0}$) under the C-D West model mortality pattern presented in col.(3) of Table 2 are **not fully** consistent with the above expectation. It is worthwhile to mention here that the mortality levels **($\mathbf{e^0_0}$)** corresponding to $_5\mathbf{LS}'_{\mathbf{a+}}$ values at ages 65, 70, and 75, which vary between 75.37 and 76.51, are remarkably close. And the values of $\mathbf{e^0_0}$ corresponding to $_5\mathbf{LS}'_{\mathbf{a+}}$ at ages 50, 55, and 60 are also sufficiently close to each other, but the estimates are relatively higher than at ages 65 and above[11]. It may be noted that the mortality levels ($\mathbf{e^0_0}$) corresponding to $_5\mathbf{LS}'_{\mathbf{a+}}$ at ages 5 to 45 exceed the highest mortality level ($\mathbf{e^0_0=80}$) of the Coale-Demeny model life table system. ***This is mainly due to the fact that the proposed technique assumes that the growth curve ($\bar{\mathbf{r}}_{\mathbf{a}}$) follows a second-degree polynomial. Whereas an empirical examination of the nature of $_5r_a$ values for Swedish female shows that the $_5r_a$ values***

between ages 5 to 50 are rather irregular and highly erratic. This suggests that $\bar{\mathbf{r}}_{\mathbf{a}}$ -curve cannot be treated as second degree between ages 5 to 50. However, the values of $_5\mathbf{r_a}$ at ages 55 and above show a systematically increasing pattern. Furthermore, at ages 55 and above the value of $\mathbf{r_{(a+5)+}}$ is sufficiently close to the average of $\mathbf{r_{a+}}$ and $\mathbf{r_{(a+10)+}}$. The above pattern of $_5\mathbf{r_a}$ **and** $\mathbf{r_{a+}}$ at ages 55 and above resembles

[11] The systematic changes in the estimates of $\mathbf{e^0_0}$ excepting for the open-ended terminal age-interval, as observed in the estimates of $\mathbf{e^0_0}$**,** presented in the coloumn (3) of the Tables 2 are primarily due to the blending error in splitting **10-cum-LSRs into 5-cum-LSRs** through the formulas **(10.1) to (10.3)** over various age-intervals. It worth noting though the effect of the error in estimating $\mathbf{h^*_{a+5}}$ [defined by $\frac{1}{2}$ **($\mathbf{C_{(a+10)+}}$ - $\mathbf{C_{a+}}$)]** as discussed under the methodology section cannot be totally ruled out, however, its impact becomes negligibly small as the magnitudes of $\mathbf{C_{a+}}$ and $\mathbf{C_{(a+10)+}}$ are sufficiently close to the mean-ages of persons aged **'a & above'** and **'a+10 & above'** respectively.

quite closely with a second degree polynomial of the growth curve ($\bar{r}_a$) beyond age 55 which is also supported
by the closeness of the estimates of $\mathbf{e^0_0}$ corresponding to $_5\mathbf{LS'_{a+}}$ values at ages 55 and above. The median mortality level ($\mathbf{e^0_0}$) out of the last five $\mathbf{e^0_0}$ values (that is, at ages 55 and above) in col.(3) of Table 2 is 76.51 years which can be taken as the final estimate of mortality level ($\mathbf{e^0_0}$) for Swedish females during 1966-70.

The values of $\mathbf{e^0_a}$ at various quinquennial ages beyond age 5 corresponding to the Coale-Demeny (1983) west model life table for females at the above median mortality level, that is, $\mathbf{e^0_0}$**= 76.51** years are sufficiently close to those of the official estimates of $\mathbf{e^0_a}$ particularly at younger ages. The above analysis indicates that the ***West model mortality pattern*** corresponding to the level $\mathbf{e^0_0}$**=76.51** resembles well with the Swedish female mortality pattern during 1966-70.

[Tables 1 & 2 to be entered here - shown at the end of the paper]

Indian Females, 1981-91

The second application refers to the enumerated age-returns of Indian females of 1981 and 1991 censuses. The relevant data are borrowed from the published reports of Indian censuses (Office of the Registrar General of India, 1987 & 1997). It is needless to emphasize that the magnitude and pattern of age misreporting in Indian censuses create considerable difficulties in estimating demographic parameters through indirect techniques. An empirical study of the values of $\mathbf{r_{a+}}$ during 1981-91 over various quinquennial ages shows a gradual-increasing trend particularly between ages 30 to 60. Irregular fluctuations seem to be more pronounced in $_5\mathbf{r_a}$ values compared to $\mathbf{r_{a+}}$ values over ages. However a regular increasing trend has been observed in the age-range 35 to 55. The above empirical investigation suggests that the growth curve ($\bar{r}_a$) for Indian females during 1981-91 resemble very closely a second-degree polynomial particularly between ages 30 to 60. Since the intercensal period is ten years, the formula (14) was applied to estimate **10-cum-LSRs (** $_{10}\mathbf{LS'_{a+}}$ **)** at ages 5 to 60 which were spitted into **5-cum-LSRs (** $_5\mathbf{LS'_{a+}}$ **)** values through the application of the formulas (15.5) to (15.7). These estimated values of ${}_5\mathbf{LS}'_{\mathbf{a+}}$ and the estimates of ${}_5\mathbf{r_a}$ **and** $\mathbf{r_{a+}}$ along with the average persons years lived during the intercensal period 1981-91 are presented in Table 3. The estimated values of ${}_5\mathbf{LS}'_{\mathbf{a+}}$ along with the corresponding mortality levels are presented in Table 4.

It has been mentioned earlier that **5-cum-LSRs** values alone cannot determine a life table uniquely unless the appropriate mortality pattern is also known. An empirical study with various model patterns it is found that the model mortality patterns in the ***UN South Asian model life tables*** work reasonably well for Indian females during the decade 1981-91. The estimates of mortality levels ($\mathbf{e_0^0}$) associated with the estimated ${}_5\mathbf{LS}'_{\mathbf{a+}}$ values during 1981-91 are presented in the col.(3) of Table 2. The estimates of mortality levels ($\mathbf{e_0^0}$) associated with the estimated ${}_5\mathbf{LS}'_{\mathbf{a+}}$ values are sufficiently closed in the age-range 30 to 65. These findings indicate that the assumptions involved in estimating **10-cum-LSRs** (${}_{10}\mathbf{LS}'_{\mathbf{a+}}$) from the corresponding **10-cum-CSRs** (${}_{10}\mathbf{CS}'_{\mathbf{a}}$) are largely satisfied and the errors in age reporting get diluted considerably due to dampening effects of cumulation. Thus, the median $\mathbf{e_0^0}$ value (**56.667 years**) of the $\mathbf{e_0^0}$ values corresponding to the estimated (observed) ${}_5\mathbf{LS}'_{\mathbf{a+}}$ values at ages 30 to 65 under ***UN South Asian Model life table system*** may be taken as the appropriate mortality level for the Indian females during the decade 1981-91. The life tables at ages 5 and above under UN South Asian model mortality corresponding to the life expectancy of 56.667 may be taken as the appropriate adult mortality tables consistent with the set of ${}_5\mathbf{LS}'_{\mathbf{a+}}$ values at ages 5 and above for Indian females during 1981-91. The **e(x)** values at ages 5 and above corresponding to the median mortality level (that is, **e(0)=56.667 during 1981-91**) which are obtained by linear interpolation are presented in col.(4) of the Table 2. These estimated **e(x) values** during 1981-91 are also sufficiently close to those obtained by SRS during 1981-91 as the average values of **e(x)** during the period 1981-85, and 1986-90 obtained by SRS. These values are shown in col.(5) of Table 4.

It has been shown in an earlier study (Lahiri and Meneges, 2004) that the mortality pattern for Indian females during 1971-81 is closer to C-D South Model Life Table System, whereas that for the period 1981-91, as indicated by the present study, resembles closely with the UN South Asian Model Life Table System. The above findings are also supported by some other studies on mortality pattern carried out on the basis of SRS data on age-specific death rates during the periods 1971-81, and 1981-91 respectively (Roy, and Lahiri, 1987; and Lahiri, and Rao and Srinivasan, 2003).

[Tables 3 & 4 to be here - shown at the end of the paper}

Summary and Conclusions

An attempt has been made in this paper in developing an analytical relationship between '10-year ***conventional*** census survival ratios' **(10-CSRs)** and its associated '10-year *conventional* life table survival ratios' **(10-LSRs**), defined by the ratio $_5L_{x+10} / _5L_x$ in life table terminology, in any closed population that follows a ***generalized population model*** applicable to any population. It has been shown that the equality between **10-CSRs** and **10-LSRs,** which is assumed in ***conventional census survivorship approach in assessing intercensal mortality***, holds good ***only when the population under study is either stationary or stable***. An attempt has also been made to develop a formula for estimating **10-LSRs** from two census enumerations having intercensal period ***other than ten years not necessarily multiple of 5***.

Using the property of ***cumulative*** census survivorship ratios, based on ***cumulated*** census age-returns beyond certain quinquennial ages, in controlling the effects of age-misreporting to a large extent, a formula has also been proposed in this study for estimating '10-year ***cumulative*** life table survival ratios' **(10-cum-LSRs)**, defined by the ratio T_{x+10} / T_x in life table terminology, from the corresponding **'10-year *cumulative*** census survivorship ratios'**(10-cum-CSRs).** The development of such a relationship has been carried out in this study under the assumption of generalized model of age-structure of a closed population having decennial population censuses. Furthermore while developing the requisite formulas for estimating **10-LSRs, and 10-cum-LSRs** from two enumerations as mentioned above, it is assumed that the age-specific growth curve follows a second-degree polynomial. A formula has also been derived here in estimating **10-cum-LSRs** from two consecutive census age-returns of a population irrespective of width of the interval between the two consecutive censuses, ***not necessarily multiple of 5***. It is worthwhile to mention here that the equality between the **10-cum-CSRs and 10-cum-LSRs** holds good ***only when the population under study is stationary***.

There exists an ***exact formula (mathematical identity)***, shown in this paper, through which the above set of **10-cum-LSRs (along with the knowledge of 10-LSRs)** can be split into **5-cum-LSRs** (defined by the ratios of the form, T_{x+5} / T_x) which, in turn, can be translated into **5-LSRs**, the 5-year ***conventional*** life table survival ratios through another identity. However, in this study for splitting **10-cum-LSRs** into **5-cum-LSRs** we have used another approximate formula which is not only operationally more convenient than the exact formula but also helps in smoothing the fluctuations in **10-cum-LSRs** estimated from the census age-returns. Even in the presence of age misreporting the values of **10-cum-CSR** over ages usually show a systematically declining pattern of positive fractions, as one would normally expect in contrast to the erratic pattern of **10-CSRs** due to age misreporting. Thus one would expect that the values of **5-cum-LSR** over ages estimated through the above procedure would also show similar regular pattern in contrast to those of **5-LSRs**. Further investigations may be carried out in obtaining adjusted series of **5-LSRs** by smoothening the ***latent irregularities*** in **5-cum-LSRs** due to age-misreporting through a suitable mathematical graduation formula, such as -- Brass-type two-parameter logit model (see, Lahiri, 1983 for further discussions). Such an attempt would be particularly helpful in obtaining smooth series of **5-LSRs** ***without smoothening of the distorted age-data and/or the raw (conventional) census survival ratios due to age misreporting.*** **It is well recognized that the** ***smoothing of defective age-data or survival ratios are rather arbitrary in nature and often influenced by personal predilections.***

The formulas developed here for estimating longevities at various quinquennial ages have been applied to the age-data of Sweden, well known for its accuracy, followed by those of Indian data that are heavily distorted due to age misreporting. The analyses indicate that the technique proposed here works quit well and the estimates of life expectancies at various quinquennial age beyond age five are sufficiently close to those of the official estimates based on age-specific death rates for periods of the study populations considered here.

REFERENCES

Bennett, N.G., and Horiuchi, S. (1981) **Estimating the completeness of death registration in a closed population**, ***Population Index***, **47(2): 207-221**.

Bhat, P.N. Mari (1987) ***Mortality in India: Levels, Trends, and Patterns***, Ph.D. Dissertation, **University of Pennsylvania (Michigan, University Microfilms, 1987).**

--------------- (1995) **Age misreporting and its impact on adult mortality estimates in south Asia**, ***Demography-India*****, 24(1): 59-90.**

Coale, A.J. (1984) **Life table construction on the basis of two enumerations of a closed population**, ***Population Index*****, 50(2); 193-213.**

Coale, Anslay J., Paul Demeny, and Barbara Vaughan (1983) **Regional Model Life Tables and Stable Populations**, Second Editions, Academic Press, New York.

Lahiri, Subrata (1983) ***Life Table Construction from Population Age-Distributions Suffering from Response Biases in Age-Reporting: A New Technique (not requiring age-smoothing) with Application to Indian Census age-returns***, **Unpublished Ph.D. Dissertation, Indian Statistical Institute, Calcutta.**

--------------- (1985) **Life table construction from two enumerations subject to age-misreporting: A new technique not requiring age-smoothing**, In ***Dynamics of Population and Family Welfare 1985,(edited by K. Srinivasan, and S.Mukerji)***, **Himalaya Publishing House, Mumbai, 1985: 73-101.**

--------------- (1990) **Some new approaches to the estimation of life expectancies at older ages**, In ***Dynamics of Population and Family Welfare, 1989***, (eds. by Srinivasan and K.B. Pathak), pp.315-341.

--------------- (2002) **A Study of Mortality in the Indian Subcontinent and Some Selected Developed Countries: Perspectives, Contrasts and Challenges.** ***Presented at the Meeting at the Airlie Conference Center, Warrenton, Virginia, USA during November 1-5, 2002 under the Fulbright New Century Scholars Program, 2001-2002.***

Lahiri, Subrata and Lysander Menezes (2004) **Estimation of Adult Mortality from Two Enumerations of a Destabilized Population Subject to Response Biases in Age-Reporting,** An earlier version was presented at the PAA 2000 Annual meeting at Los Angeles during March 22-25, 2000. The revised and updated version of the paper, carried out by the first author, was included as one of the part of the research project report (unpublished) entitled, "**A Study of Mortality in India and Some Selected Developed and Developing Countries: Perspectives, Contrast and Challenges**" **under the** ***Fulbright New Century Scholars Award*** **to the first author during 2001-2002.**

Subrata Lahiri, Arni S. R. Rao, and S. Srinivasan (2003) "**Role of Age-Specific Growth Rates on** Population Ageing in Some Developed and Developing Countries -- A Comparative **Study"**, A Revised Version of the paper presented in an IUSSP sponsored ***Workshop on Age Structural Transitions and Policy Implications*** held at Phuket, Thailand, Nov., 2000.

Office of the Registrar General of India (1988) ***Social and Cultural Tables (Tables C-1 to C-6)***, Part IV-A, Census of India, 1981, Registrar General and Census Commissioner of India, New Delhi.

Office of the Registrar General of India (1997): ***Social and Cultural Tables, Part IV A- C Series, Vol.2, Census of India 1991,*** Registrar General and Census Commissioner of India, New Delhi

Palloni, A., and R. Kominski (1984) **Estimation of adult mortality using forward and backward projection**, ***Population Studies***, 38(3):479-93.

Preston, S.H., and A.J. Coale (1982) **Age structure, growth, attrition, and accession: A new synthesis, *Population Index*, 48(2):217-259.**

Preston, S.H., and N.G. Bennett (1983) **A census-based method for estimating adult mortality, *Population Studies*, 37(1): 91-104**

Preston, Samuel H., and Subrata Lahiri (1991): **A short-cut method for estimating death registration completeness in destabilized populations**, ***Mathematical Population Studies (U.S.A.)***, Vol. 3, No.1, pp.39-51.

Roy, T. K. and Subrata Lahiri **(1988) "Recent Levels and Trends in Mortality in India and Its Major States: An Analysis Based on SRS Data"**, In ***Dynamics of Population and Family Welfare 1987***, Himalaya Publishing House, 1988, pp.279-349 (ed. K. Srinivasan, and S. Mukerji).

Sweden, Statistika Centrallyrån (1974) ***Livslangdstabeller for Artiondet 1961-1970 [Life Tables for the Decade 1961-1970]***, Stockholm.

United Nations (1967) ***Methods of Estimating Basic Demographic Measures from Incomplete Data***, ***U.N. Manual IV*** (Population Studies No.42: ST/SOA/SER.A/42). **Department of Social Affairs, New York.**

United Nations (1983) **Indirect Techniques for Demographic Estimations, Manual X**, Population Studies No.81, Department International Economic and Social Affairs, (ST/ESA/SER.A/81).

Technical Appendix-A

Estimation of Mean-age Beyond age a (C_{a+})

The mean-age of persons aged 'a & above' is generally obtained as the weighted average of the mid-points of various 5-year age-intervals together with an arbitrarily fixed mean-age of the open-ended terminal age-interval taking the respective population size in various age-intervals including the terminal open-ended age-interval as the weights. To avoid the ambiguity of arbitrarily fixing mean age of the terminal open-ended age-interval, Preston & Lahiri (1991) proposed the following procedure. The mean-age of elderly persons (that is the terminal open-ended age group) aged 'x and above', where x assumes one of the following values -- 65, 70, 75 or 80, which can be obtained through the successive application of the following formulas **under the assumption of sectional stability of the population aged 'x & above'** (Preston, and Lahiri, 1991).

$$A_{D_{x+}} = \frac{\ln b_x - \ln(b_x - r_{x+})}{r_{x+}} - \frac{r_{x+}}{2}\sigma^2(D_{x+}), \text{ ------ (A1) and } A_{P_{x+}} = \frac{1-(b_x - r_{x+})A_{D_{x+}}}{r_{x+}} \text{ -------(A2)}$$

where, $$b_x = \frac{1}{\overline{N}(x+)}\left[\frac{{}_5\overline{N}_{x-5} * {}_5\overline{N}_x \ln({}_5\overline{N}_{x-5}/{}_5\overline{N}_x)}{5*({}_5\overline{N}_{x-5} - {}_5\overline{N}_x)}\right] \text{ ------(A3), and hence } C_{x+} = A_{P_{x+}} + x \text{ -----(A4)}$$

The notations used in the above formulas are as follows:

A_{Px+} : Mean age of persons aged 'x & above' measured from age x;

A_{Dx+} : Mean age at deaths of persons aged '**x & above**' measured from age **x;**

b_x : 'Birth-day rate at age x', defined by the ratio $\overline{N}(x)/\overline{N}(x+)$ where $\overline{N}(x)$ denotes the number of persons at exact age **x**;

r_{x+} **:** Exponential rate of growth of persons aged '**x & above**'; and

$\sigma^2(D_{x+})$ **:** Variance of the age-distribution of deaths at age '**x and above**'.

In their paper, Preston & Lahiri (1991) proposed to use some standard values of $\sigma^2(D_{x+})$ based on Coale-Demeny (1983) stable population under "West" model mortality pattern. The above formula (A3) for estimating $\overline{N}(x)$ and hence $b_x = \overline{N}_x / \overline{N}(x+)$ rests on the assumption that the population at older ages '65 & above' is approximately stable and the death rate at exact age x - $\mu(x)$ - remains almost constant in the decennial age-range (x-5, x+5) (see, **Appendix-B).**

Technical Appendix-B

An Approximation for Estimating $\overline{N}(x)$, the Number of Persons at Exact Age x, from the Quinquennial Age Data Assuming Local Stability in the 10-year Age-interval (x-5, x+5)

Let us consider a population that grows exponentially with constant rate of growth **r(a) = r** and constant death rate, that is **μ(a) =μ ,** for all **'a'** in the ten-year age-interval **(x-5, x+5)**. Under the assumption of local stability of the population along with above assumptions regarding growth rate and death rate in the age-interval, the number of persons at exact age 'a' can be expressed by the following equation:

$$\mathbf{N(a) = N(x-5) * e^{-k(a-x+5)}}\text{, where } \mathbf{k = r+\mu}\text{, and } \mathbf{x-5 \le a \le x+5} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(B1)}$$

Now, integrating both sides of the **(B1)** within the domains of integration **x-5 to x**, and **x to x+5** we get ${}_5\overline{N}_{x-5}$ and ${}_5\overline{N}_x$, the number of persons in the age-intervals (x-5, x) and (x, x+5) respectively as follows:

$${}_5\overline{N}_{x-5} = \frac{\overline{N}(x-5) * (1-e^{-5k})}{k} ------\text{(B2)}$$

and

$${}_5\overline{N}_x = \frac{\overline{N}(x-5) * e^{-5k} * (1-e^{-5k})}{k} ---\text{(B3)}$$

$$\text{or } {}_5\overline{N}_x = \frac{\overline{N}(x) * (1-e^{-5k})}{k} --------\text{(B3.1)}$$

Since $\overline{\mathbf{N}}\mathbf{(x) = \overline{N}(x-5) * e^{-5k}}$, which follows from the equation (B1). Now, dividing **(B2)** by **(B3)** we get

$${}_5\overline{N}_{x-5} / {}_5\overline{N}_x = e^{5k} ---------------\text{(B4)}$$

Now, using **(B3.1)** and **(B4)** and simplifying the results together with the knowledge of the parameter $\mathbf{k = \ln\left({}_5\overline{N}_{x-5} / {}_5\overline{N}_x\right)}$ that follows from the equation **(B4)**, we finally get the desired formula for $\overline{\mathbf{N}}\mathbf{(x)}$ as given below:

$$\overline{N}(x) = \frac{{}_5\overline{N}_{x-5} * {}_5\overline{N}_x * \ln\left({}_5\overline{N}_{x-5} / {}_5\overline{N}_x\right)}{5 * \left({}_5\overline{N}_{x-5} - {}_5\overline{N}_x\right)} ---\text{(B5)}$$

Table 1
Estimation of 10-cum-LSRs ($_{10}LS'_{a+}$) at various quinquennial ages starting with age 5 from two enumerations for Swedish females during 1966-70.

Age	Av.No. of persons-years during 1966-70[1]	Av. annual exponential rate of growth during 1966-70	Av. annual rate of growth of person aged a and above[2]	Mean age of person aged a and above[3]	10-cum-LSRs at age a[4]
a	$_5N_a$	$_5r_a$	r_{a+}	C_{a+}	$_{10}LS'_{a+}$
(1)	(2)	(3)	(4)	(5)	(6)
5	267,024	0.015746	0.005022	40.5313	0.8768676
10	259,265	- 0.005044	0.004182	43.1110	0.8652897
15	285,619	- 0.033251	0.004942	45.6227	0.8519670
20	309,826	0.003975	0.008737	48.4174	0.8401491
25	267,775	0.048595	0.009313	51.5488	0.8297285
30	226,859	0.014550	0.004732	54.3529	0.8189904
35	226,394	- 0.016020	0.003656	56.7482	0.7953834
40	249,171	- 0.025007	0.006073	59.1123	0.7643704
45	263,172	- 0.001657	0.010931	61.7089	0.7255138
50	260,557	0.001220	0.013420	64.5185	0.6778498
55	254,801	0.000380	0.016390	67.4441	0.6194971
60	235,205	0.016034	0.021392	70.5508	0.5429924
65	201,225	0.020365	0.023563	73.8134	0.4467501
70	160,846	0.022252	0.025260	77.1641	0.3297182
75	114,792	0.025615	0.027477	80.6007	
80	66,452	0.033675	0.029542	84.0393	
85+	37,059	0.022130	0.022130		

---

[1] Borrowed from a paper by Preston and Bennett (1983)

2) $$\hat{r}_{a+} = \frac{\sum_{x=a}^{85+} {}_5\hat{N}_x * {}_5\hat{r}_x}{\sum_{x=a}^{85+} {}_5\hat{N}_x};$$

[3] The procedures for estimating C_{a+}'s including C_{80+} are described in the text (see, the section - ***Estimation of mean age beyond age a (C_{a+}) and h^*_a***.

[4] Obtained through the formula (15) applicable for any intercensal interval, not necessarily multiple of 5.

Table 2

Estimation of adult mortality for Swedish Females, 1966-70 consistent with the estimated 5-cum-LSRs during 1966-70.

Age	Estimated 5-cum.LSR at age a [1]	C-D West Model Mortality Pattern		Official life expectancy at age a[2]
		Mort. level - M-L (e^0_0) w.r. to 5-cum-LSR at age a in col.(2)	Est. life expectancy at age a based on the selected life table[3]	
a	${}_5LS'_{a+}$	$ML(e^0_0)$	$e^0_a(E)$	$e^0_a(0)$
(1)	(2)	(3)	(4)	(5)
5	0.9395289	**	72.52	72.52
10	0.9333056	**	67.58	67.64
15	0.9266081	**	62.63	62.71
20	0.9198029	**	57.71	57.83
25	0.9137410	**	52.80	52.96
30	0.9079328	**	47.92	48.10
35	0.8983878	**	43.06	43.27
40	0.8830191	**	38.24	38.49
45	0.8629535	**	33.50	33.80
50	0.8374228	79.350	28.87	29.21
55	0.8049950	78.330	24.39	24.75
60	0.7615672	77.493	20.10	20.43
65	0.7018020	76.513	16.04	16.35
70	0.6195151	75.372	12.39	12.58
75	0.5322198	76.468	9.25	9.32

Note: **: **The estimated mortality level exceeds the highest life expectancy in the Coale-Demeny model life table system.**

[1] **5-cum-LSRs** are obtained from **10-cum-LSRs** (shown in col.(6) of Table 1) through the formulas (15.6), (15.7) & (15.8)

[2] Borrowed from the Swedish female life table for the period 1965-70 prepared by the National Central Bureau

[3] Based on the model life table corresponding to the median mortality level (i.e., $\mathbf{e^0_0=77.003}$ years) among the mortality levels ($\mathbf{e^0_0}$) in col.(3) with respect to the estimated ${}_5LS'_{a+}$ values at ages 50 to 75 shown in col. 2 of Statistics, Sweden (Sweden, Statistiska Centralbyrån, 1974).

Table 3

Estimation of 10-cum-LSRs ($_{10}LS'_{a+}$) at various quinquennial ages starting with age 5 from two enumerations, 10-years apart for Indian females during 1981-91.

Age	Av.No. of persons-years during 1981-91[1]	Av. annual exponential rate of growth during 1981-91[2]	Av. annual rate of growth of person aged a and above[3]	Mean age of person aged a and above[4]	10-cum-LSRs at age a[5]
a	$_5N_a$	$_5r_a$	r_{a+}	C_{a+}	$_{10}LS'_{a+}$
(1)	(2)	(3)	(4)	(5)	(6)
5	49,652,788	0.017570653	0.023219597	28.6620195	0.85391735
10	43,734,707	0.014473409	0.024277694	32.6220127	0.85860905
15	33,431,337	0.020562952	0.026216794	36.5931346	0.84987598
20	32,541,487	0.027054384	0.027222531	39.9852172	0.81876718
25	29,642,523	0.033380320	0.027257682	43.6411829	0.78957088
30	24,506,168	0.031941041	0.025819925	47.4387642	0.75609841
35	21,824,228	0.027510732	0.024344449	51.0460792	0.72317697
40	17,920,929	0.020408831	0.023477844	54.7571797	0.69144307
45	15,503,167	0.021920637	0.024369208	58.3149566	0.64942836
50	12,894,648	0.020756286	0.025190589	61.9411378	0.62399416
55	9,186,421	0.029004201	0.026907375	65.5922837	0.52435897
60	9,800,594	0.021568551	0.026110670	68.6696703	0.45359785
65	5,516,200	0.030378269	0.029216141	72.8816028	
70	4,504,559	0.023024005	0.028491471	76.2398295	
75+	4,335,224	0.034187554	0.034187554		

1) $_5\hat{N}_a = \frac{_5P_a^{91} - {_5P_a^{81}}}{10.\ _5\hat{r}_a}$, 2) $_5\hat{r}_a = 0.1 * \ln(_5P_a^{91} / {_5P_a^{81}})$, 3) $\hat{r}_{a+} = 0.1 * \ln(P_{a+}^{91}/P_{a+}^{81})$.

where the enumerated age-data of 1981 and 1991 censuses were borrowed from respective census publications of India.

4). The procedures for estimating C_{a+}'s including C_{65+} are described in the text--
Estimation of mean age beyond age a (C_{a+}) and h^*_{a+}; and

5). Obtained through the formula (14) applicable for decennial censuses.

Table 4

Estimation of adult mortality for Indian females, 1981-91 consistent with the estimated 5-cum-LSRs during 1981-91.

Age	Estimated 5-cum.LSR at age a [1]	U.N South Asian Model Mortality Pattern		Official life expectancy at age a[2]
		Mort. level - M-L (e^0_0) w.r. to 5-cum-LSR at age a in col.(2)	Est. life expectancy at age a based on the selected life table[3]	
a	${}_5LS'_{a+}$	$ML(e^0_0)$	$e^0_a(E)$	$e^0_a(0)$
(1)	(2)	(3)	(4)	(5)
5	0.92281153	59.42673249	62.961	62.125
10	0.92534317	71.99357935	58.750	58.240
15	0.92424637	>e(0)=75	54.112	53.740
20	0.91333267	72.87685444	49.579	49.440
25	0.89668083	63.92552302	45.095	45.165
30	0.87900706	57.80945382	40.616	40.900
35	0.85991540	54.18003794	36.182	36.560
40	0.84091143	54.20752298	31.795	32.210
45	0.81860034	55.40918900	27.486	27.945
50	0.79786230	60.12846095	23.298	23.240
55	0.75631429	58.54276515	19.375	19.885
60	0.69835287	55.52476405	15.793	16.260
65	0.64952529	59.27795497	12.649	13.220
70			9.920	10.555

[1] The value of **5 cum-LSRs (${}_5LS'_{a+}$)** were obtained from **10-cum-LSRs (${}_{10}LS'_{a+}$)**, shown in col.(6) of Table 3, through the formulas **(15.6), (15.7) & (15.8).**

[2] Based on the model life table corresponding to the $\mathbf{e^0_0=56.667}$, the median mortality level out of the mortality levels ($\mathbf{e^0_0}$) in col.(3) with respect to the estimated ${}_5LS'_{a+}$ values at ages 30 to 65 of col.(2);

[3] Average of the life expectancies estimated by SRS for the period 1981-85, and 1986-90.