DEMOGRAPHIC RESEARCH

Research Article

**On stochastic comparisons of population
densities and life expectancies**

Maxim Finkelstein

Table of Contents

On stochastic comparisons of population densities and life expectancies

Maxim Finkelstein [1]

Abstract

Cohort measures, describing a lifetime random variable are easily and unambiguously obtained using standard tools. On the contrary, the lifetime random variable, and therefore life expectancy, for the period setting cannot be unambiguously defined without additional simplifying assumptions. For non-stationary populations the corresponding conventional period measures should be justified in some way. Our paper is based on Bongaarts and Feeney (2002). We consider different measures of life expectancy and compare them for specific populations using stochastic ordering of the corresponding random variables. This gives possibility to look at the problem in a more general way.

[1] Department of Mathematical Statistics, University of the Free State PO Box 339, 9300 Bloemfontein, Republic of South Africa,
E-mail: FinkelM.SCI@mail.uovs.ac.za and Max Planck Institute for Demographic Research, Rostock, Germany

1. Introduction

Mortality is declining with time, which is a consequence of better conditions of life. By "conditions of life" or mortality conditions we mean the whole range of factors with health care being probably the major one. As a result, life expectancy at birth is increasing. For instance, Oeppen and Vaupel (2002) state that female life expectancy in the country with maximum life expectancy is increasing every year for approximately 3 months.

'True' mortality rates and life expectancies for birth cohorts can be derived simply and unambiguously using standard statistical tools as the corresponding lifetime random variable has a proper statistical meaning in this case. Period life tables and age-specific mortality rates have been used for many years for obtaining life expectancy and other demographic measures. However, changing in time mortality or, more generally, population non-stationarity makes life much more complicated. The 'statistical' origin of this complexity is in the fact that unlike the cohort setting, the lifetime random variable for the period setting cannot be unambiguously defined without additional simplifying assumptions. Under rather stringent "proportionality" conditions and for a specific population Bongaarts and Feeney (2002) suggested a candidate for the 'true' life expectancy and obtained analytical relations for the deviation from the conventional measure in this case. They showed that it is quite substantial and varies from 1.5 to 3 years for different developed countries. Therefore, when mortality rates are decreasing with time, the conventional approach gives larger values of life expectancy, as compared with this measure. The attractive and at the same time restrictive feature of this model is that it fixes mortality conditions at a given instant of time and assumes that a specifically defined population exists in this fixed environment.

As the notion of a 'true' life expectancy is somehow elusive, it is more a matter of definitions and assumptions. The average age at death, e.g., might be also a widely used measure of life expectancy, if it were more easily calculated in practice.

This is a technical note with a main goal to suggest some types of stochastic ordering for populations, to explore some possibilities in defining life expectancy and to compare mathematically these definitions using suggested stochastic ordering of populations. In our next publication on the topic we intend to present further results in this direction and to illustrate them by the data.

2. Stochastic comparison of population densities

Denote by $N(x,t), x \geq 0$ a population density (age-specific population size) at time t - a number of persons of age x. See Keiding (1990) and Arthur and Vaupel (1984) for discussion of this quantity. It can be also interpreted as the continuous population surface over age x and time t.

Let X_t denote a random age at time t of a person who is picked out at random (with equal chances) from a population of size $\int_0^\infty N(u,t)du$, where for convenience of notation the upper limit of integration is ∞ (instead of the maximal attainable age ω). Therefore we interpret X_t as a random age in a population with a density $N(x,t), x \geq 0$. Let

$$f(x,t) = \frac{N(x,t)}{\int_0^\infty N(x,t)dx} \tag{1}$$

and

$$F(x,t) = \Pr(X_t \leq x) = \frac{\int_0^x N(u,t)du}{\int_0^\infty N(u,t)du} \tag{2}$$

define the probability density function (pdf) and the cumulative distribution function (Cdf) of X_t, respectively. Therefore $F(x,t)$ is the age Cdf which corresponds to the population density $N(x,t), x \geq 0$. It can be equivalently interpreted as a proportion of persons in our population whose age does not exceed x. Note, that equations (1) and (2), in fact, define for a given population the estimates of the pdf and the Cdf, respectively (observed period values). Distribution $F(x,t)$ can be formally written via some hypothetic rate $\tilde{\mu}(x,t)$ by a standard exponential representation:

$$F(x,t) = 1 - \exp\left\{-\int_0^x \tilde{\mu}(u,t)du\right\}, \tag{3}$$

where: $\tilde{\mu}(x,t) = \dfrac{f(x,t)}{1-F(x,t)} = \dfrac{N(x,t)}{\displaystyle\int_x^\infty N(u,t)du}$.

The mean age of a population with a population density $N(x,t), x \geq 0$ is:

$$e(t) \equiv \frac{\displaystyle\int_0^\infty xN(x,t)du}{\displaystyle\int_0^\infty N(x,t)du} = \int_0^\infty \exp\left\{-\int_0^x \tilde{\mu}(u,t)du\right\}dx . \tag{4}$$

The mean age of a population is not sensitive to the tail of the Cdf $F(x,t)$, which also could be of interest in demographic studies. The simplest tail measure of an age structure is its $q-$ quantile:

$$q_F = F_t^{-1}(q), \tag{5}$$

where $F_t^{-1}(\cdot)$ is an inverse function of $F(x,t)$ for the fixed t. Therefore, the proportion of persons in a given population with age exceeding q_F is q. For high quantile values ($q \geq 0.8$) q_F indicates the start of an old age zone of a population. It is worth mentioning that the mean age of a population is quite different from life expectancy. In the latest U.S. life table, life expectancy is 77.3 years and the implied stationary population has the mean age around 40.4 years.

Let $N(x,t), x \geq 0$ and $N^*(x,t), x \geq 0$ be the densities for two populations with random ages X_t and X_t^*, respectively. Specific types of these population densities will be considered in the next section. Above all, these specific populations are needed for dealing with 'proper' random variables in defining life expectancy.

Stochastic comparison of X_t and X_t^* is of interest. We say that the age X_t^* is stochastically larger (Ross, 1996; Shaked and Shantikhumar, 1993) than the age X_t and write

$$X_t^* >_{st} X_t , \tag{6}$$

if the corresponding age distribution functions are ordered as

$$F^*(x,t) < F(x,t), \forall x > 0 .\tag{7}$$

Inequality (7) is equivalent to:

$$\frac{\int_0^x N^*(u,t)du}{\int_0^\infty N^*(u,t)du} < \frac{\int_0^x N(u,t)du}{\int_0^\infty N(u,t)du}; \forall x > 0 \quad ,\tag{8}$$

and the population density $N^*(x,t), x \geq 0$ gives larger probabilities to ages beyond x, than $N(x,t), x \geq 0$. It also follows from (7) and definitions (4) and (5) that the corresponding mean population ages and the quantiles are ordered in a similar way, respectively: $e^*(t) > e(t)$, $q_{F^*} > q_F; 0 < q < 1$, which illustrates the fact that stochastic ordering (6) is stronger than comparison in the mean values and comparison of quantiles (Finkelstein, 2003).

Stochastic comparison of the same population at different time instants can be also of interest. Inequality

$$X_{t_2} >_{st} X_{t_1}; t_2 > t_1\tag{9}$$

means that the population with age structure $N(x,t_2), x \geq 0$ is stochastically older than population with age structure $N(x,t_1), x \geq 0$, which certainly is the case in practice (under reasonable restrictions on fertility and migration), because mortality rates are declining with time. If this inequality holds for all ordered t_1 and t_2 in some interval of time, we say that the population is aging in this interval of time. In fact, this is a definition of an *aging in a strong sense* population. The corresponding ordering of the mean values defines *aging in a weak sense.*

3. Specific populations modeling

Let $\mu(x,t)$ denote the mortality rate as a function of age x and time t for a population with a density $N(x,t), x \geq 0$:

$$\mu(x,t) = \lim_{\delta \to 0} \frac{(N(x+\delta,t+\delta) - N(x,t))/d\delta}{N(x,t)} \qquad (10)$$

As, given a population, this rate is observed, there is nothing wrong with it in a period setting. On the other hand, a lifetime cumulative distribution function cannot be unambiguously defined only via $\mu(x,t)$ without additional assumptions.

Consider a population, which is closed to migration and experience a constant number of births B annually. The population density $N(x,t), x \geq 0$ in this case can be defined via the corresponding cohort survival function:

$$N(x,t) = B \exp\left\{ -\int_0^x \mu(u,t-x+u)du \right\} = B l_c(x,t-x) , \qquad (11)$$

where $l_c(x,t-x)$ denote the survival probability of a cohort of age x, which was born at time $t-x$.

Let $N(x,t), x \geq 0$ be the same as in (11) observed population density. As in Bongaarts and Feeney (2002), we 'freeze' now mortality conditions for a fixed time t:

$$N(x,t) = B \exp\left\{ -\int_0^x \mu^*(u,t)du \right\}. \qquad (12)$$

The function $\mu^*(x,t)$ can be interpreted as the mortality rate for a stationary population with the density $N(x,t), x \geq 0$. Therefore the corresponding lifetime random variable is defined via $\mu^*(x,t)$ in a usual way. Note, that although the integrals (and therefore the corresponding survival functions) in equations (11) and (12) are obviously equal, the integrands are not equal. The corresponding exponential representation via the mortality rate $\mu(x,t)$ for the same density reads (Preston and Coale, 1982):

$$N(x,t) = B \exp\left\{ -\int_0^x \mu(u,t)du \right\} \exp\left\{ -\int_0^x \rho(u,t)du \right\}, \qquad (13)$$

where $\rho(u,t)$ is the intensity of population growth: $\rho(x,t) = \dfrac{\partial N(x,t)/\partial t}{N(x,t)}$.

It can be easily seen from equations (11), (12) and (13) that

$$\int_0^x \rho(u,t)du = \int_0^x (\mu(u,t-x+u)-\mu(u,t))du$$

and that $\rho(x,t) = \mu^*(x,t)-\mu(x,t)$ in this specific case (see also Arthur and Vaupel (1984)). Equation (13) can be *formally* transformed to:

$$N(x,t) = B\exp\left\{-\int_0^x \mu(u,t)D(u,t)du\right\}, \tag{14}$$

where

$$D(x,t) = 1 + \frac{\rho(x,t)}{\mu(x,t)} \tag{15}$$

is a factor, which adjusts, in accordance with equation (11), the period mortality rate at a fixed time instant t to the mortality rate with a changing time variable. If, e.g., a population is growing: $D(x,t) > 1$. Under additional assumptions (see later), Bongaarts and Feeney (2002) show that with a suitable in practice accuracy $D(x,t)$ *does not depend on age* x.

Consider now a hypothetical population (also closed to migration and with a constant birth rate) and define a hypothetical population density $N^*(x,t), x \geq 0$ via the mortality rate $\mu(x,t)$:

$$N^*(x,t) = B^*\exp\left\{-\int_0^x \mu(u,t)du\right\}. \tag{16}$$

Therefore, $\mu(x,t)$ can be also interpreted as the mortality rate for the stationary population with the population density $N^*(x,t), x \geq 0$.

Expressions (11)-(16) will be used for comparing Cdfs of X_t^* and X_t and also for comparing different definitions of life expectancy.

Lemma 1. *Let* $f(x)$ *and* $g(x)$ *be continuous functions such that* $g(x)$ *is decreasing and* $\int_0^\infty f(x)dx < \infty$. *Then*

$$\frac{\int_0^x f(u)g(u)du}{\int_0^\infty f(u)g(u)du} > \frac{\int_0^x f(u)du}{\int_0^\infty f(u)du}, \forall x > 0.$$

Proof. Applying the mean value theorem:

$$\frac{\int_0^x f(u)g(u)du}{\int_0^\infty f(u)g(u)du} = \frac{\int_0^x f(u)g(u)du}{\int_0^x f(u)g(u)du + \int_x^\infty f(u)g(u)}$$

$$= \frac{g(0,x)\int_0^x f(u)du}{g(0,x)\int_0^x f(u)du + g(x,\infty)\int_x^\infty f(u)du} > \frac{\int_0^x f(u)du}{\int_0^\infty f(u)du},$$

where $g(0,x)$ and $g(0,\infty)$ are the corresponding mean values, which exist due to our assumptions. As $g(x)$ is decreasing: $g(0,x) > g(0,\infty)$, and, therefore, the inequality follows.

The following result shows that random ages of two populations under consideration are ordered as in definition (6). In demographic terms it states that:

If mortality is declining, the observed population will be *stochastically younger* than that of the stationary population implied by $\mu(x,t)$.

Proposition 1. *Let mortality rate* $\mu(x,t)$ *decrease in time t. Assume that population densities* $N(x,t), x \geq 0$ *and* $N^*(x,t), x \geq 0$ *are given by equations (13) and (16), respectively. Then stochastic ordering (6) ((7)) holds.*

Proof. In accordance with inequality (8) and equations (13) and (16), we must show that:

$$\frac{\int\limits_0^x \exp\left\{-\int\limits_0^y \mu(u,t)du\right\}dy}{\int\limits_0^\infty \exp\left\{-\int\limits_0^y \mu(u,t)du\right\}dy} < \frac{\int\limits_0^x \exp\left\{-\int\limits_0^y \mu(u,t)du\right\}\exp\left\{-\int\limits_0^y \rho(u,t)du\right\}dy}{\int\limits_0^\infty \exp\left\{-\int\limits_0^y \mu(u,t)du\right\}\exp\left\{-\int\limits_0^y \rho(u,t)du\right\}dy}.$$

As mortality rate decreases with time, the population is growing. Therefore, $\rho(u,t) > 0$ and the corresponding exponential function in the integrand is monotonically decreasing with y. The result immediately follows from Lemma 1 after

noting that $\int\limits_0^\infty \exp\left\{-\int\limits_0^y \mu(u,t)du\right\}dy < \infty$.

The useful for stochastic comparison of populations alternative proof of Proposition 1, based on a totally different concept, is deferred to the Appendix.

Remark. Proposition 1 can be generalized to the case when mortality rates for each fixed x are decreasing in average, which means that the following integral is decreasing in t for $\forall x > 0$:

$$\frac{\int\limits_0^t \mu(x,u)du}{t}.$$

A rather cumbersome mathematical proof of this fact is out of the scope of this paper.

The following result (Lemma 2) can be interpreted in demographic terms as:

If mortality is declining, the age at death in the actual population is *stochastically younger* than that of the corresponding stationary population.

We feel that this result can be also generalized to the case of declining in average mortality rates, but this should be still proved.

Lemma 2. *Let mortality rate $\mu(x,t)$ decrease in time t. Assume that population densities $N(x,t), x \geq 0$ and $N^*(x,t), x \geq 0$ are given by equations (13) and (16), respectively.* Then:

$$\frac{\int\limits_0^x \mu(u,t)N^*(u,t)du}{\int\limits_0^\infty \mu(u,t)N^*(u,t)du} < \frac{\int\limits_0^x \mu(u,t)N(u,t)du}{\int\limits_0^\infty \mu(u,t)N(u,t)du} ; \ \forall x > 0 .$$

Proof. Substituting relations (11) and (16) into this inequality:

$$\frac{\int\limits_0^x \mu(y,t)\exp\left\{-\int\limits_0^y \mu(u,t)du\right\}dy}{\int\limits_0^\infty \mu(y,t)\exp\left\{-\int\limits_0^y \mu(u,t)du\right\}dy} < \frac{\int\limits_0^x \mu(y,t)\exp\left\{-\int\limits_0^y \mu(u,t)du\right\}\exp\left\{-\int\limits_0^y \rho(u,t)du\right\}dy}{\int\limits_0^\infty \mu(y,t)\exp\left\{-\int\limits_0^y \mu(u,t)du\right\}\exp\left\{-\int\limits_0^y \rho(u,t)du\right\}dy}$$

which is proved using exactly the same argument as while proving Proposition 1 after noting that

$$\int\limits_0^\infty \mu(y,t)\exp\left\{-\int\limits_0^y \mu(u,t)du\right\}dy = 1 < \infty .$$

4. Definitions of life expectancy

4.1 Period life expectancy at birth (conventional)

Period life expectancy at birth at time t is usually defined as the average age at death that would be observed for a group of persons (a cohort) who experience, over the course of their lives the age specific mortality rates $\mu(x,t)$ observed during the time period:

$$e(0,t) = \int\limits_0^\infty l(x,t)dx . \tag{17}$$

where $l(x,t)$ is a life table survival probability:

$$l(x,t) = \exp\left\{-\int_0^x \mu(u,t)du\right\}.$$ (18)

When mortality rates are changing over time, as it was mentioned, the corresponding lifetime random variable cannot be unambiguously defined and other approaches for obtaining life expectancy can be used. Bongaarts and Feeney (2002) have suggested the following model:

4.2 Life expectancy at current mortality conditions

Definition (17) is based only on age-specific mortality rates obtained without any direct information on the population age structure. Life expectancy at current mortality conditions is defined by Bongaarts and Feeney (2002) for the specific 'frozen' population, described by equation (12) as

$$e^*(0,t) = \int_0^\infty l^*(x,t)dx,$$ (19)

where $l^*(x,t)$ is a survival function, defined via the mortality rate $\mu^*(x,t)$ similar to equation (18).

Alternatively, $e^*(0,t)$ can be written via the observed mortality rate as an integral of the survival function in equation (11):

$$e^*(0,t) = \int_0^\infty \left(\exp\left\{ -\int_0^x \mu(u,t-x+u)du \right\} \right) dx = \int_0^\infty l_c(x,t-x)dx,$$

which is, in fact, the cross-sectional average life (CAL), introduced in Gulliot (2003).

4.3 Average age at death

Another period measure, which is used for calculating life expectancy at least under the assumption of a constant in time birth rate, is $A(t)$ -the average age at death at current mortality rates:

$$A(t) = \frac{\int\limits_0^\infty x\mu(x,t)N(x,t)dx}{\int\limits_0^\infty \mu(x,t)N(x,t)dx} \ . \tag{20}$$

This measure already reflects population characteristics and can be estimated from the period life table data, although, as it was mentioned, the calculations are not so easy compared with $e(0,t)$.

Under additional assumptions (see the next section) Bongaarts and Feeney (2002) showed that $e^*(0,t) = A(t)$. Note that we do not need these additional assumptions for comparing $e(0,t)$ with $A(t)$ in the next section.

In stationary populations all three measures are equal:

$e(0,t) = e^*(0,t) = A(t)$.

When mortality is changing with time, however, they are different and the use of conventional $e(0,t)$ should be justified in some way.

Some other 'corrected' measures based on $e^*(0,t)$ (e.g., weighting in some suitable sense the inputs of different cohorts) can be also considered.

4.4 The mean population age

It is clear that the mean population age $e(t)$, defined by equation (4), is not a very good indicator of life expectancy, but it can be used as some helpful additional characteristic. Assume for simplicity, that mortality rate $\mu(x,t)$ is increasing in age x for all $x \geq 0$. The following comparison of $e(t)$ with the average age at death $A(t)$ is straightforward:

$$\frac{\int\limits_0^\infty x\mu(x,t)N(x,t)dx}{\int\limits_0^\infty \mu(x,t)N(x,t)dx} > \frac{\int\limits_0^\infty xN(x,t)dx}{\int\limits_0^\infty N(x,t)dx} \Rightarrow A(t) > e(t) \ . \tag{21}$$

Indeed, the 'modified' population density is $\mu(x,t)N(x,t); \, x \geq 0$. It is clear that under the assumption of increasing in age mortality rate $\mu(x,t)$ this density gives larger probabilities to ages beyond x, than $N(x,t), x \geq 0$, which results in inequality similar to inequality (8) and, finally, in inequality (21).

5. Comparing life expectancies

5.1 Comparison of $e(0,t)$ with $e^*(0,t)$

We shall make this comparison for a population that experiences no migration and constant number of births annually. In fact, $e^*(0,t)$ was defined in Section 4 only for this specific population type. As population is growing (declining mortality rates):

$$\mu^*(x,t) - \mu(x,t) > 0; \, \forall x \geq 0 , \tag{22}$$

which obviously leads to the corresponding ordering of life expectancies (see equations (17) and (19)) and to a difference $\Delta(t)$:

$$\Delta(t) \equiv e(0,t) - e^*(0,t) > 0. \tag{23}$$

This is a general result for the population of the defined type, which can be also formulated as:

$$\Delta(t) = \int_0^\infty \left(\exp\left\{ -\int_0^x \mu(u,t)du \right\} \right) dx - \int_0^\infty \left(\exp\left\{ -\int_0^x \mu(u,t-x+u)du \right\} \right) dx .$$

Bongaarts and Feeney (2002) make additional assumptions for estimating $\Delta(t)$, which they call the *tempo bias*. They assume that changes in the population density $N(x,t), x \geq 0$ due to mortality decline are modelled as an *age-independent* shift $s(t)$ to higher ages:

$$\begin{aligned} N(x,t) &= B, \quad x < s(t) \\ N(x,t) &= N(x - s(t),0), \quad x \geq s(t) \end{aligned} \tag{24}$$

Equations (24) lead to the same shift in mortality rates. Formally this is a rather stringent assumption, although assuming the Gompertz law for mortality curves with fixed t, we immediately arrive at the following Gompertz shift model, as an exponential function 'converts shifts into multipliers':

$$\mu^*(x,t) = \mu^*(0,t)\exp(bx) , \qquad (25)$$

where $b > 0$ is a parameter. This relationship was verified in Bongaarts and Feeney (2002) for a number of countries with remarkably good accuracy. Therefore, the contemporary demographic data supports assumption (24) as well. It was proved by these authors that:

$$\mu(x,t) = \left(1 - \frac{de^*(0,t)}{dt}\right)\mu^*(x,t) . \qquad (26)$$

Equation (26) shows that when life expectancy $e^*(0,t)$ is increasing the observed mortality rate $\mu(x,t)$ is smaller than the non-observed $\mu^*(x,t)$. Using numerical procedure Bongaarts and Feeney (2002) obtained the values of $e^*(0,t)$ and the corresponding tempo bias (23). It turned out that the average tempo bias, e.g., for females in France, Japan, Sweden and USA for the period 1980 to 1995 is rather large: 2.3 years, 3.3 years, 1.6 years, and 1.6 years, respectively. However, a question still remains: is $e^*(0,t)$, defined for a specific population under rather stringent conditions, the best candidate for the 'true' life expectancy?

5.2 Comparison of $e(0,t)$ with $A(t)$

The following lemma is a direct consequence of definition (16):

Lemma 3. *Let $N^*(x,t), x \geq 0$ be a density for a hypothetical population defined by equation (16). Then the average age at death for this population $A^*(t)$ is equal to conventional life expectancy $e(0,t)$:*

$$A^*(t) \equiv \frac{\int_0^\infty x\mu(x,t)N^*(x,t)dx}{\int_0^\infty \mu(x,t)N^*(x,t)dx} = e(0,t) \qquad (27)$$

Proposition 2. *Let mortality rate $\mu(x,t)$ decrease in time t. Assume that population densities $N(x,t)$ and $N^*(x,t)$ are given by equations (13) and (16), respectively. Then conventional life expectancy $e(0,t)$ is larger then the average age at death*

$$e(0,t) - A(t) > 0 . \qquad (28)$$

Proof. In accordance with Lemma 3 and definition (20), we must prove that

$$\frac{\int_0^\infty x\mu(x,t)N^*(x,t)dx}{\int_0^\infty \mu(x,t)N^*(x,t)dx} - \frac{\int_0^\infty x\mu(x,t)N(x,t)dx}{\int_0^\infty \mu(x,t)N(x,t)dx} > 0 \qquad (29)$$

As ordering of survival functions trivially leads to the same ordering of the corresponding mean values, inequality (29) immediately follows from Lemma 2 (the sign of inequality of this lemma will be opposite for survival functions). Alternatively, averaging procedures (Vaupel, 1992) can be also used for proving Proposition 2.

As it was mentioned, for proving inequality (28) we do not need additional proportionality assumptions. Note, that an obvious from general considerations (under given assumptions) Proposition 2 is a simple corollary of stochastic orderings of Section 3 that are, in fact, the main mathematical results of this paper.

5.3 Comparison with a hypothetical cohort

The following alternative comparison with a hypothetical cohort, which allows for time-dependent birth rates, can be also helpful:

Let W denote the maximum age in our life table, e.g., $W = 110$ years. The age structure $N(x,t), x \geq 0$ means that $B(t-x)$ persons were born at $t-x$, $N(x,t)$ of which had survived up to t (no migration). Let us shift the 'life trajectories' of

survivors backwards on $W - x$ units of time. This means that the whole population $\int_{0}^{W} N(u,t)du$ will be born at $t - W$ and the cohort for the whole population can be considered. As mortality rates are declining with t :

$$\mu(x, t - W + x) > \mu(x, t) \tag{30}$$

This inequality also means that $e(0,t) > e_s(0,t)$, where $e_s(0,t)$ denotes life expectancy of the described shifted cohort.

6. Concluding remarks

Life expectancy in reliability and survival analysis studies is the simplest characteristic of a lifetime distribution, as it usually has a clear cohort-type meaning. In demography, however, the period life tables are the main data source and due to population non-stationarity (e.g., the declining in time mortality rates) there can be different definitions of life expectancy based on period life tables. What is the 'true life expectancy' is still an open-ended question. In Section 5 we compare the conventional life expectancy with two other possible measures.

Bongaarts and Feeney (2002) considered a specific, closed to migration and with constant number of births annually population with frozen mortality conditions. They showed that when mortality rates are declining in time, the conventional observed life expectancy substantially overestimates the life expectancy for this model, obtained using numerical methods.

In the current paper we prove analytically that conventional life expectancy also majorizes the average age at death. We consider stochastic comparisons of age distributions for specific populations and obtain inequalities for the corresponding mean values as a simple consequence. We feel that stochastic comparisons of populations should be explored to a greater extent in the future studies.

7. Acknowledgements

References

Arthur, W.B., and Vaupel, J.W. (1984). Some general relationships in population dynamics. *Population Index,* 50, 214-226.

Bongaarts, J., and Feeney G. (2002). How long do we live? *Population and Development Review*, 28, 13-29

Finkelstein, M.S. (2001). A note on stochastically acceptable quality (2001). Quality and Reliability Engineering International, 17, 19-24.

Finkelstein, M.S. (2003). On some measures and distances for positive random variables. Applied Stochastic Models in Business and Industry, 19, 133-146.

Guillot, M. (2003). The cross-sectional average length of life (CAL). Population Studies, 57, 41-54.

Keiding, N. (1990). Statistical inference in the Lexis diagram. Philosophical Transactions of the Royal Society of London, A 332, 487-509.

Oeppen J., and Vaupel J.W. (2002). Broken limits to life expectancy. *Science*, 296, 1029-1031.

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

Ross S.M. *Stochastic Processes*. John Wiley & Sons, 1996.

Shaked, M., and Shantikhumar J. (1993). *Stochastic Orders and Their Applications*. Academic Press, Boston.

Vaupel, J.W. (1992). Analysis of Population Changes and Differences. Unpublished book.

Appendix

The following simple probabilistic reasoning can help in interpretation of ordering (8). Denote by $N_G(x), x \geq 0$ an age structure of a stationary population with age Cdf $G(x)$:

$$G(x) = \frac{\int_0^x N_G(u)du}{\int_0^\infty N_G(u)du}.$$

Let $x_1, x_2, ..., x_{n-1}$ be $n-1$ be the points dividing population into n equal parts of size d, which represent number of persons in each interval $[x_{k-1}, x_k], k = 1,2,...,n; x_0 = 0, x_n = \infty$:

$$\int_0^{x_1} N_G(u)du = \int_{x_1}^{x_2} N_G(u)du = ... = \int_{x_{n-1}}^\infty N_G(u)du = d.$$

Thus:

$$G(x_k) = \frac{k}{n}, k \leq n.$$

Denote, as previously, by $G(x,t)$ the age Cdf for a 'real population' with declining mortality rate, closed to migration and with constant birth rate. This population immediately starts to be stationary when mortality rate stops changing. Let $t = t_d$ be some fixed instant of time: the 'demographic regime' for $t < t_d$ is more severe than for $t = t_d$. Let, specifically, $G(x)$ be defined by a constant in time demographic regime, which is experienced by the 'real population' at $t = t_d$. As mortality rates are declining with time:

$$G(x_k, t_d) = \frac{kd - \sum_1^k d_i}{nd - \sum_1^n d_i}, k \leq n$$

where d_i is the number of persons that were lost (died) in the corresponding interval of time due to the change of the constant regime model to the one with improving regime. It is clear that d_i increase with the number of the interval:

$$d_1 < d_2 < ... < d_n.$$

Given this ordering:

$$\frac{kd - \sum_1^k d_i}{nd - \sum_1^n d_i} < \frac{k}{n}, k \le n$$

which is a discrete analogue of (8) By letting $n \to \infty$, the continuous analogue of this inequality (which corresponds to inequality (8)) can be obtained.