



*Demographic Research* a free, expedited, online journal  
of peer-reviewed research and commentary  
in the population sciences published by the  
Max Planck Institute for Demographic Research  
Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY  
[www.demographic-research.org](http://www.demographic-research.org)

, citation and similar papers at [core.ac.uk](http://core.ac.uk)

brought to

provided by Research P

## DEMOGRAPHIC RESEARCH

**VOLUME 24, ARTICLE 22, PAGES 527-550  
PUBLISHED 31 MARCH 2011**

<http://www.demographic-research.org/Volumes/Vol24/22/>

DOI: 10.4054/DemRes.2011.24.22

*Research Article*

### **Significance of life table estimates for small populations: Simulation-based study of standard errors**

**Sergei Scherbov**

**Dalkhat Ediev**

© 2011 *Sergei Scherbov & Dalkhat Ediev.*

*This open-access work is published under the terms of the Creative Commons Attribution NonCommercial License 2.0 Germany, which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit.*

*See <http://creativecommons.org/licenses/by-nc/2.0/de/>*

## Table of Contents

1	Introduction: Data and methods	528
2	Preliminary formal considerations and simulations design	529
3	Life table procedures. Imputations for the open age interval	535
4	Results	536
4.1	General overview	536
4.2	Standard errors of life expectancy estimates	537
4.3	Biases	539
4.4	Normality of life expectancy estimate's distribution	540
5	Illustrative examples	541
6	General recommendations	543
7	Acknowledgements	545
	References	546
	Appendix: Supplementary tables	547

## **Significance of life table estimates for small populations: Simulation-based study of standard errors**

**Sergei Scherbov<sup>1</sup>**

**Dalkhat Ediev<sup>2</sup>**

### **Abstract**

We study bias, standard errors, and distributions of characteristics of life tables for small populations. Theoretical considerations and simulations show that statistical efficiency of different methods is, above all, affected by the population size. Yet it is also significantly affected by the life table construction method and by a population's age composition. Study results are presented in the form of ready-to-use tables and relations, which may be useful in assessing the significance of estimates and differences in life expectancy across time and space for the territories with a small population size, when standard errors of life expectancy estimates may be high.

---

<sup>1</sup> Vienna Institute of Demography, Austrian Academy of Sciences. World Population Program, International Institute for Applied Systems Analysis. E-mail: [Sergei.Scherbov@oeaw.ac.at](mailto:Sergei.Scherbov@oeaw.ac.at).

<sup>2</sup> Vienna Institute of Demography, Austrian Academy of Sciences. E-mail: [Dalkhat.Ediev@oeaw.ac.at](mailto:Dalkhat.Ediev@oeaw.ac.at).

## 1. Introduction: Data and methods

Life expectancy is a key characteristic of human longevity and development, and policies worldwide aim to increase it. While effective policies can be based on informative monitoring systems, estimating life expectancy for small populations may run into difficulties because of the small number of events and insufficient exposures, which lead to uncertainty in estimating death rates. This makes the development of methodology for estimating and comparing life expectancy for small populations a high priority.

Using the Monte Carlo simulation approach, Silcocks, Jenner, and Reza (2001), Toson, Baker, and the Office of National Statistics (2003), Eayres and Williams (2004), as well as Williams et al. (2005) evaluated methodologies for the estimation of small-area life expectancy in the United Kingdom (UK) context. They showed that life expectancy at birth is distributed normally and estimates of its standard error are distributed with a significant skew for the small population size. They also demonstrated that traditional life table methodology without special corrections for age bands with zero deaths in a small population performs quite well, and that the choice of the minimum age of the open age interval and modeling the mortality in that interval are important for estimating life expectancy and its standard error. Based on the simulated dependency of standard errors on population size, a minimum population years-at-risk size of 5,000 for estimating life expectancy at birth was recommended in the UK context. However, the age composition of a small population in all the tests was fixed and was only scaled up and down depending on the simulated population size. Apart from that, the effects of life expectancy level on estimation accuracy were not explored. In this paper we extend previous research by including the effects of population age composition and life expectancy level.

Our work extends the previous research in several directions. First we confirm some of the findings in the literature in a wider context of mortality schedules and population structures. We conduct simulations based on all available male and female life tables for Austria, Italy, Japan, Spain, Sweden, and the UK, which were chosen as being representative of the variety of mortality situations in currently low-mortality countries. We use data from the *Human Mortality Database* (2010). For each life table scrutinized, we consider five stable population age compositions corresponding to -2%, -1%, 0%, 1%, and 2% annual population growth rates. Based on those mortality and population schedules, we consider eight population sizes of 1,000, 5,000, 10,000, 25,000, 50,000, 100,000, 250,000 and 1 million people (in total, 43,680 populations).

Second we present the empirical relations between the standard error of life expectancy indicators and the corresponding life table and population characteristics (life expectancy, and population growth rate). Third we evaluate standard errors for

both the life expectancy at birth and the life expectancy at age 60—two measurements that are essential in the context of policies oriented toward population aging and pension systems. Fourth we provide a more in-depth analysis of the normality of life expectancy estimates for small populations, and illustrate that age composition can crucially affect the normality of estimates (which is important for establishing confidence limits and the significance of the variation observed in terms of life expectancy).

In contrast to previous work, we consider indicators of unabridged life tables. We also study the estimates for abridged life table calculations based on the age groups 0, 1, 5, 10, ..., 85+ years. However both the previous work and our own study (not reported here) indicate that using abridged as opposed to unabridged life tables has only a small effect on estimation accuracy compared with the procedure chosen for the open age interval. At the same time, we find that estimates for abridged life tables tend to be systematically biased when the age composition deviates from that of the stationary population, irrespective of the population size. (These distortions are caused by a deviation from the stationary age composition within individual age intervals.) Therefore avoiding the use of abridged life tables is recommended unless the age composition of the population is fairly close to stationary.

In the three works cited above, the open age interval was chosen to start at 85, 90, or 95 years; if no deaths occurred in the open age interval, the corresponding mortality rate was taken from a known life table and not from the simulated population. We use a different approach, adjusting the open age interval in such a way that there is at least one death in it; hence, we do not use (unavailable in practice) rates from a theoretical life table in order to infer a life table for the simulated population.

In this paper we present auxiliary formal relations and describe how the simulations are designed; we then continue with a discussion of life table calculations, followed by the presentation of results. We conclude by presenting illustrative case studies and general recommendations. The paper is supplemented by an Appendix with tabular material.

## **2. Preliminary formal considerations and simulations design**

Our study is based on simulations, as analytical approximations of standard errors of life expectancy estimates (Chiang 1984) are biased for small population (Eayres and Williams 2004). However formal considerations are still useful for understanding the mechanism behind the standard errors and biases of life expectancy estimates. Those considerations are in this section.

## The model

Individuals in a population of size  $N_x$  at exact age  $x$  all have a probability of dying  $q_x$  during one year. There is no migration. The number of deaths is binomially distributed.

A maximum likelihood estimator of  $q_x$  is  $\hat{q}_x = \frac{D_x}{N_x}$  where  $D_x$  is the observed number of deaths (Chiang 1984, eqs. (5.1) and (5.12)). The variance of  $\hat{q}_x$  equals  $\frac{q_x(1-q_x)}{N_x}$ , which can be estimated as  $\frac{\hat{q}_x(1-\hat{q}_x)}{N_x} = \frac{\hat{q}_x^2(1-\hat{q}_x)}{D_x}$  (Chiang 1984, eqs. (3.7) and (3.8)).

Assuming a constant force of mortality  $m_x$ , the death probability  $q_x$  equals  $1 - \exp(-m_x)$ . The occurrence-exposure rate  $\hat{m}_x = \frac{D_x}{P_x}$  ( $P_x$  stands for the population person-years exposed at age  $x$  to  $x+1$ , which may be approximated by the mid-year population for all age groups except the youngest and oldest ones) is a maximum likelihood estimator of  $m_x$ . It is unbiased, consistent, and asymptotically normal when the number of deaths ( $D_x$ ) is large (Rao 1973, Chiang 1984). The asymptotic variance of  $\hat{m}_x$  equals  $\frac{m_x(1-q_x)}{P_x}$  (Chiang 1984, eq. (3.5)). Using the constant force of mortality assumption and the invariance properties of maximum likelihood, one finds that  $\hat{q}_x = 1 - \exp(-\hat{m}_x)$ .

The model described above is applied to each of a number of age intervals  $(x, x+1)$ , with age  $x$  running from 0 to 109, while  $x=110+$  represents the open interval for the highest ages. Independence across ages is assumed.

## Why a bias in life expectancy estimates?

We begin by illustrating why and how life expectancies can be biased when they are estimated from small population data.

In his classic monograph, Chiang (1984:161) shows that sample life expectancy is “an unbiased estimate of the corresponding unknown true expectation of life.” This conclusion was based on the assumption that “the observed expectation of life [life expectancy calculated from observed death rates—Scherbov and Ediev] at a given age is the sample mean lifetime of individuals living beyond this age.” Yet, such an

assumption can only be asserted for cohort life tables obtained from observations over individual lifetimes. In our study we focus on estimates of period life expectancies; therefore, Chiang's proposition about non-bias in the life expectancy may be violated (and, indeed, is, as follows from our and others' simulations and the following formal relations).

Consider first the estimate of probability of surviving to a given age, which is a function of accumulated mortality rates  $\hat{m}_x$  (these represent the mortality rates and not the individual lifetimes upon which the period life table calculations are based):

$$\hat{l}_x \approx e^{-\sum_{y=0}^{x-1} \hat{m}_y} \tag{1}$$

Mortality rates are unbiased estimates of the underlying theoretical rates:  $E(\hat{m}_x) = m_x$  (Chiang 1984). Separating the expected values in (1) and expanding by Taylor's theorem, yields:

$$\hat{l}_x \approx e^{-\sum_{y=0}^{x-1} m_y} e^{-\sum_{y=0}^{x-1} (\hat{m}_y - m_y)} = l_x e^{-\sum_{y=0}^{x-1} (\hat{m}_y - m_y)} \approx l_x \left( 1 - \sum_{y=0}^{x-1} (\hat{m}_y - m_y) + \frac{1}{2} \left[ \sum_{y=0}^{x-1} (\hat{m}_y - m_y) \right]^2 \right). \tag{2}$$

Hence, assuming independence and non-bias in individual mortality rates, the estimated survival probability is biased upwards:

$$\begin{aligned} E(\hat{l}_x) &\approx l_x - l_x \sum_{y=0}^{x-1} E(\hat{m}_y - m_y) + \frac{1}{2} l_x E \left[ \left[ \sum_{y=0}^{x-1} (\hat{m}_y - m_y) \right]^2 \right] = \\ &= l_x + \frac{1}{2} l_x E \left[ \sum_{y=0}^{x-1} (\hat{m}_y - m_y)^2 \right] = l_x + l_x \frac{1}{2} \sum_{y=0}^{x-1} \sigma^2(\hat{m}_y) \end{aligned} \tag{3}$$

Similar relations apply to the probabilities of surviving from a given age  $x$  to another given age  $a$ . The life expectancy is the sum of such survival probabilities, which explains why it must be biased upwards. The results of adding more terms into the Taylor series expression in (3) cannot be exactly calculated analytically. However, at small population sizes the third central moment of binomial distribution is negative (at typically low  $\hat{m}$  mortality levels, the occurrence-exposure rate  $\hat{m}_x = \frac{D_x}{P_x}$  is approximately

proportional to the binomially distributed enumerator). Therefore, both third- and fourth-degree terms in the Taylor expansion yield additional upward bias.

### How are biases related to population size?

As noted above, variance of the occurrence-exposure mortality rate is inversely proportional to the population exposed  $P_x$  in the respective age group; asymptotically (Chiang 1984),

$$\sigma^2(\hat{m}_x) = \frac{m_x(1-q_x)}{P_x}. \quad (4)$$

Hence, bias in (3) and in life expectancies must increase as population size decreases. eq. (4) also suggests that as population size increases, standard errors of estimates of life expectancy decrease (asymptotically) as an inverse square root of population size.

### How are standard errors related to population size?

To roughly estimate variance of the survival probability, we drop the quadratic term in (2):

$$\sigma^2(\hat{l}_x) \approx \sigma^2 \left[ l_x \left( 1 - \sum_{y=0}^{x-1} (\hat{m}_y - m_y) \right) \right] = l_x^2 \sum_{y=0}^{x-1} \sigma^2(\hat{m}_y), \quad (5)$$

which, given eq. (4), implies that the standard errors of survival probabilities are asymptotically inversely proportional to the square root of the population size. The same applies to the probabilities of surviving from one given age to another and to life expectancy. At small population sizes when the contribution of the third and fourth moments in the Taylor series become considerable, there is an additional increase in standard errors. This effect is also visible in simulations. In our study the inverse proportionality between the standard errors and the population size may be used for populations of at least 5,000 people.



### How skewed are the distributions of the estimates?

The relation of survival probability to mortality rates is also indicative of how skewed its distribution might be. At typically low mortality levels, the occurrence-exposure rate

$\hat{m}_x = \frac{D_x}{P_x}$  is approximately proportional to the binomially distributed and not skewed

enumerator. A full Taylor's expansion in (2) would combine non-skewed distributions generated by odd-powered summands and positively skewed distributions generated by even-powered summands. Hence, survival probabilities (and thus life expectancies) must be positively skewed. This means, in particular, that at sufficiently small population numbers, the distribution of life expectancies will deviate from normal.

### On Chiang's approximation

Chiang (1984:161-165) proposed a useful method for approximating standard errors of life table estimates for small areas. His recurrent method was based on first-order approximation to Taylor's series of life expectancy as a function of survival probabilities. Earlier studies (Toson, Baker, and the Office of National Statistics 2003; Eayres and Williams 2004) suggested, in the UK context, the effectiveness of Chiang's approximate method. Eayres and Williams report a good fit of the method for the standard error of life expectancy at birth at large population sizes; yet, at population size 5,000 the reported bias of the method already amounts to a decimal digit. We also studied the method using Japanese female life tables in 1947, 1977, and 2007 and came to results similar to those reported earlier. We found that the method yields strong biases at small population size (in our simulations, the method underestimates, on average, the standard error of life expectancy at birth by up to 0.3 years at population size 5,000 and by up to 0.8 years at population size 1,000; those biases are up to 0.4 and 6 years, respectively, for the life expectancy at age 60). At large population size, its bias, being small in absolute value, amounts to 5% of the true standard error (results are similar for life expectancy both at birth and at age 60). A drawback, in the context of our study, of Chiang's method for the standard errors is its inability to provide sample distributions of the estimated life expectancies and their standard errors. Therefore, our prime method of studying the standard errors in the work was based on simulations, and not on Chiang's approximation. Although Chiang's method may be used without significant problems starting from a population size of about 10,000, we also provide ready-to-use tables, which might be more convenient in practice at any population size.

## Simulation design

Each life table defines a stationary population (e.g., Keyfitz 1977). For one particular life table, we selected a certain population size  $N$ , and simulated populations with sizes  $N$  equal to 1,000, 5,000, 10,000, 25,000, 50,000, 100,000, 250,000 and 1 million people. The life table defines a probability of dying  $q_x$  at each age interval. Given  $N$ , we also know the number of people  $N_x$  at each age  $x$ . The number of deaths in each age interval was drawn from a binomial distribution with probability  $q_x$  and size  $N_x$ . One simulation run resulted in one specific value for the number of deaths in each age interval. This resulted in one life table and one set of values for the life expectancies at various ages. Repeated simulation gave us many sets of such life expectancies, and we report below the average values and standard errors of  $e_0$  and  $e_{60}$  across all simulations.

In addition to the case of a stationary population, we also simulated life tables based on stable populations with growth rates  $r$  equal to -2%, -1%, +1%, and +2%. For a given life table and a given growth rate  $r$ , the age structure of the corresponding stable population can be constructed (e.g., Keyfitz 1977). Given the size  $N$  of each stable population, we computed  $N_x$ , the number of people at each age  $x$ , and simulated life tables and life expectancy values as described above for the case of a stationary population.

## How many simulations per sample?

The number of simulations used in our study (25,000) is considerably higher than that used in the previous literature (2,000 by Silcocks, Jenner, and Reza 2001 as well as Toson, Baker, and the Office of National Statistics 2003; 10,000 by Eares and Williams 2005). Such a high number was chosen so as to reduce statistical errors of the outcome of the simulations to an acceptable minimum, as described next. The standard error of normal sample standard error  $S$  is given as  $\sigma_s \approx \sigma \frac{1}{\sqrt{2(n-1)}}$ , where  $\sigma$  is an unknown standard deviation estimated by  $S$ , and  $n$  is the sample size (Ahn and Fessler 2003). At  $n = 2000$ , the standard error amounts to about 1.6% of the standard deviation, which, being relatively small, may nonetheless considerably affect the outcome of the estimates (especially given the need to study the normality of the estimates and their confidence limits). We increased the number of simulations to 25,000 so that the relative standard error of the standard error falls below 0.5%.

### **3. Life table procedures. Imputations for the open age interval**

Small population size creates specific problems when a life table is being constructed in the usual way (see details in Eayres and Williams 2004). In particular, the absence of deaths at the open age interval implies immortality. Toson, Baker, and the Office of National Statistics (2003) as well as Eayres and Williams (2004) showed that life tables with zero death rates at age groups other than the open age group perform better than those with artificially imputed low death rates (we also came to a similar conclusion based on simulations of the Russian case, not presented here).

For the open age interval, Toson Baker, and the Office of National Statistics (2003) and Eayres and Williams (2004) proposed to impute an externally determined mortality (e.g., from the national life table) for the open age interval with no deaths observed. We have examined this method and, indeed, extra knowledge about mortality at open age intervals improves the life expectancy estimates considerably. However in many practical cases there is no basis for assuming that old-age mortality in a certain small population will be exactly the same as that observed elsewhere or on a nationwide basis. Often the very purpose of estimating life expectancy for small areas is to reveal the differences; for this purpose imputing standard mortality at open age intervals may not be sufficient.

We therefore present here an alternative approach, where the boundary of the open age interval is lowered to such a level (from the original level of 110 years) that it comprises at least one observation of death. As rough as it may be, this method performed better in our simulations than alternatives with a minimum of 2, 3, ..., 7 death observations in the open age interval (we do not present the results for those alternatives here). Except for very small and growing populations, standard errors of life expectancy estimates produced by this method were comparable to standard errors of estimates produced by imputing the theoretical mortality from the original life table for the open age interval. For a stationary population of 1,000 people, the former standard error is about 20% higher than the latter; for 2,000 people it is 5% higher; for 5,000 people, 2% higher; and for a stationary population of 25,000 people, 1% higher.

## 4. Results

### 4.1 General overview

In this section we outline the general variation of estimation biases and standard errors according to population size, stable growth rates, and mortality levels. A more detailed analysis of the factors of standard errors follows in the sections below.

Although there are distinguishable differences in the results for males and females, the differences are far smaller than the standard errors themselves. We therefore pool all the results together, irrespective of the gender of the population.

Mortality level and population age structure, on the other hand, have strong effects on the outcomes. Standard errors for estimates of life expectancy at birth tend to peak at life expectancy at birth of around 50 years, while standard errors of life expectancy at age 60 increase monotonically as life expectancy at birth increases. Population growth increases standard errors for life expectancy at age 60, while its effect on standard errors for life expectancy at birth interacts with the level of the life expectancy.

The effect of the population size, the most important driver of standard errors, may be modeled as a square-root function, as suggested by the theoretical considerations above. Starting from population size 5,000, the standard errors rescaled to populations of 1,000 people—

$$s_0 = SD_0 \sqrt{N/1000} , \quad (6)$$

$$s'_{60} = SD_{60} \sqrt{N/1000} , \quad (7)$$

$$s_{60} = SD_{60} \sqrt{N_{60+}/1000} , \quad (8)$$

—are already fairly constant, where  $SD_0$  and  $SD_{60}$  are the standard errors for life expectancy at birth and at age 60, respectively;  $N$  and  $N_{60+}$  are the total population size and population at age 60 and above. However with extremely small population sizes, the rescaled standard errors shift upwards. This problem is particularly strong for growing populations. We therefore present results separately for populations of 1,000 and 5,000 or more people.

## 4.2 Standard errors of life expectancy estimates

Simulated standard errors for life expectancy estimates at different levels of life expectancy at birth, population size, and growth rate are presented in Appendix, Table A1. In the table we present rescaled errors (6)–(8) obtained from simulations for populations of 5,000 and over. Results for the smallest population size (1,000) are singled out because, as noted above, the square-root approximation underestimates the standard errors for populations of this size.

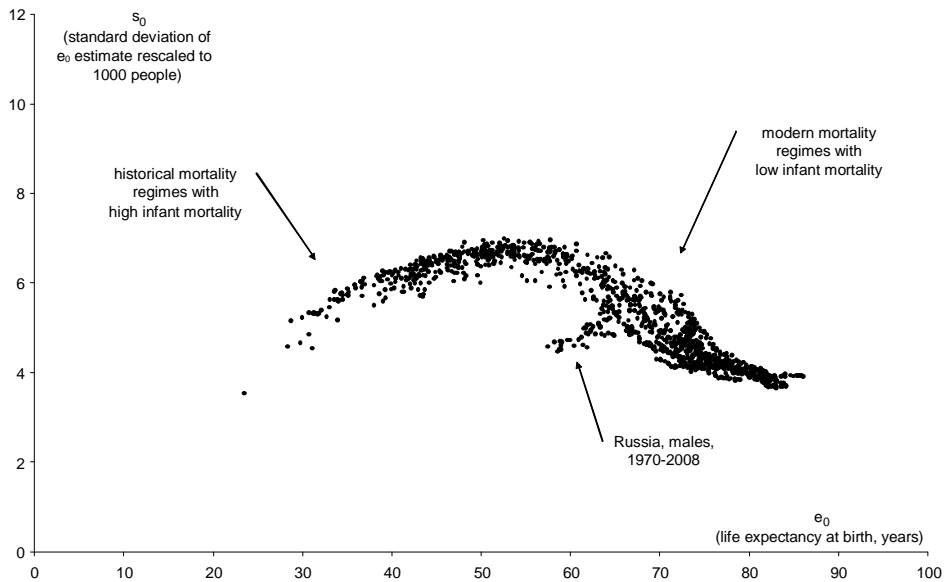
There is a curvilinear association between the rescaled standard errors  $s_0$  and the underlying true life expectancy at birth (see e.g., Figure 1 for stationary populations). This kind of association can be explained by a combination of processes with opposite effects in period mortality:

- (i) Decrease in infant and child mortality increases the role of adult mortality; this pushes up the standard errors of estimates of the life expectancy at birth. To see that, a stylized model with a known infant mortality may be considered, where  $\hat{e}_0 \approx (1 - q_0)\hat{e}_1$  and the variance  $\sigma^2(\hat{e}_0) \approx (1 - q_0)^2 \sigma^2(\hat{e}_1)$  goes up when infant mortality declines. With declining infant mortality, the effect eventually levels off.
- (ii) An opposite effect is due to mortality compression (e.g., Fries 1980). As adult mortality decreases, the distribution of period life table deaths becomes more concentrated around the mean age at death. This also suppresses the standard error of the life expectancy estimate (our simulations suggest a tight positive association between the standard error of the life expectancy estimate and the standard deviation of life table age at death).

Despite the evident overall association between the life expectancy at birth and its standard error, the particularities of mortality age patterns may strongly affect the standard errors (note the case of Russian males also presented in the figure for illustrative purposes).

Therefore, the results of our study could be used in the context of mortality estimates (both contemporary and historical) in populations with mortality resembling that observed in modern developed countries. Situations with expected deviant age patterns of mortality must be addressed separately (e.g., by conducting additional simulations).

**Figure 1: Association between standard errors of estimates and underlying theoretical values of life expectancy at birth for the stationary populations analyzed (one dot represents an average over 25,000 simulations)**



Standard errors of estimates of life expectancy at 60, not affected by the specific influence of infant mortality, follow a more consistent association with life expectancy at 60, irrespective of the population growth rate (see Figure 2 for standard errors averaged over all five population growth rates analyzed). The wide variety of simulated cases may be described by the following regression:

$$s_{60} = 0.082 \cdot e_{60} - 0.0010 \cdot e_{60}^2 + err, \quad (9)$$

with a standard error of 0.04 years. Note that the relation applies at any population growth rate; the effects of population growth rate on age structure are well captured in rescaling (8) to a population of 1,000 persons at age 60+. Also note that eq. (9) yields, naturally, more accurate estimates than those presented in Appendix Table A1, where it is the life expectancy at birth that is used as the input variable instead of the life expectancy at age 60.

**Figure 2: Association between standard errors of estimates and underlying theoretical values of life expectancy at 60 (one dot represents an average over 25,000 simulations).**

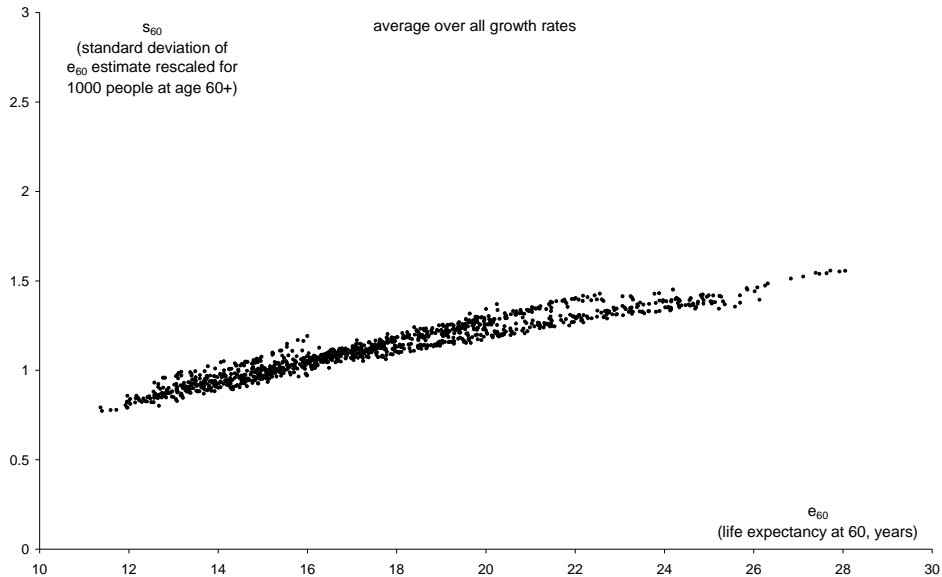


Table A1 demonstrates the importance of population age composition for accurate estimation of life expectancy in small populations. This suggests additional simulations may be required for populations whose age composition strongly differs from that of stable populations.

### 4.3 Biases

As suggested by the introductory theoretical considerations, there are upward biases in life expectancy estimates. The biases are notable for all population sizes up to 10,000 people.

For stationary and shrinking populations, the biases (for both life expectancy at birth and at age 60) amounted to about one year for populations as small as 1,000 people, and 0.2 years for populations of 5,000 people. For growing populations, these estimates must be doubled. However, the biases were significantly smaller than the

standard errors of the life expectancy estimates. According to our simulations, estimation biases may be neglected for population sizes exceeding 10,000.

Given the strong dependency of the bias on age structure, we recommend that individual corrections in each specific case should be considered depending on the actual age composition of the population at a population size of under 5,000.

Simulations indicate that there is a significant association between estimation biases and standard errors. For the sample set of stable populations examined, we found the following regression relation which may be used to roughly assess the estimation bias for life expectancy at birth:

$$Bias_0 = 0.10 \cdot SD_0 + 0.015 \cdot SD_0^2 + 0.050 \cdot SD_0 \cdot r + err \quad (10)$$

where  $\sigma_{err} = 0.1$  years ( $r$  is the stable population growth rate in percentage per annum:  $r = 1$  for 1% growth rate, etc.)

The estimation bias of  $e_{60}$  is also correlated with the standard error of the estimation:

$$Bias_{60} = 0.366 \cdot SD_{60} + 0.0265 \cdot SD_{60}^2 - 0.0094 \cdot SD_{60} \cdot r + err \quad (11)$$

with  $\sigma_{err} = 0.04$  years.

#### 4.4 Normality of life expectancy estimate's distribution

Previous research (Silcocks, Jenner, and Reza 2001; Eayres and Williams 2004; Williams et al. 2005) suggested that the distribution of estimates may be considered as approximately normal, which might simplify the practical use of standard errors of estimates (in applications such as the construction of confidence intervals, hypothesis testing, examining the significance of temporal or geographical variation of life expectancy, etc.) Strictly speaking the distribution of life expectancy estimates is not normal at any finite population size (see introductory formal considerations for explanations). For example, at 10,000 simulations, the Pearson test is powerful enough to reject normality of the simulated distribution of life expectancy even at a population of 100,000 people, when population growth is 2% per year. However, these are certain distribution percentiles rather than the normality of distributions as such, that are important for most applications. Estimates of selected percentiles derived from simulated distributions and from the corresponding normal distributions are presented in Appendix Tables A2 and A3. Percentiles obtained assuming the normality of



estimates of life expectancy at birth or at age 60 are fairly close to those obtained directly from simulated distributions at a population size of 50,000 or more. Assuming normality for a population of 5,000 or less might be discouraged, unless the tested difference in life expectancies falls far beyond the confidence limits. When studying populations of an intermediate size between 5,000 and 50,000, one must be aware of the possible effects of deviation of the age composition of the population from the stationary age composition.

## 5. Illustrative examples

Simulation results illustrate that estimations of life expectancy for small populations may be associated with quite high standard errors and biases. Those must be taken into account both in designing the system of statistical observations and in interpreting geographical, temporal, and other variations of longevity obtained from small populations. Below we present several illustrations of this kind.

### Case 1. Establishing confidence limits for life expectancy

Let life expectancy at birth be estimated at 86 years in a population of 20,000 people. What, roughly, would be the confidence limits for the actual life expectancy at the 95% confidence level, assuming stationary age composition? From Appendix Table A1, we may assess  $s_0 \approx 3.9$  (years per 1,000 persons). Hence the standard error calculated for the actual population size would be  $SD_0 = 3.9 / \sqrt{\frac{20000}{1000}} = .9$  (years). Assuming normality this yields  $e_0 = 86 \pm 1.7$  years at a 95% confidence level.

### Case 2. Examining the significance of life expectancy variation

Consider the hypothetical case of comparing life expectancy in two small populations. These populations may either represent two geographically or otherwise defined subpopulations of the total or the same population at two points in time. In the first case, we examine the significance of spatial or social variation in life expectancy, while in the second, we examine the significance of temporal variation. Suppose the two populations are characterized by the following indicators:

	Population 1	Population 2
Total population, people	20,000	50,000
Life expectancy at birth	86.0	83.5
Population at age 60 or more	5,930	15,629
Life expectancy at age 60	25.5	26.1

Then assume that the age composition of both populations is near stationary. Is the difference in life expectancy between the two populations significant (say, at the 5% significance level)?

To investigate the question above, we estimate standard errors of the estimates of life expectancy for the two populations. From Appendix Table A1 we may assess  $s_0^{(1)} \approx 3.9$  (years per 1,000 persons) for the first population and  $s_0^{(2)} \approx 3.9$  for the second population. Hence, standard errors calculated for the actual population sizes would be:

$$SD_0^{(1)} = 3.9 / \sqrt{\frac{20000}{1000}} = .9 \text{ and } SD_0^{(2)} = 3.9 / \sqrt{\frac{50000}{1000}} = .6 \text{ (years).}$$

Assuming the independence of the estimates for the two populations, we may compute the standard error of the difference between the estimates of life expectancy:

$$SD_0^{(1)-(2)} = \sqrt{(SD_0^{(1)})^2 + (SD_0^{(2)})^2} = 1.0 \text{ (years).}$$

Given the standard error and assuming normal distribution, the observed difference of 86.0-83.5=2.5 years yields  $p$ -value 1.6% (double-sided alternative) that is, *the difference is significant* at the 5% significance level. The two populations are different with respect to life expectancy at birth at the 95% confidence level.

Let us examine the significance of the difference in life expectancy at age 60. From Appendix Table A1 we obtain  $s_{60}^{(1)} \approx 1.5$  and  $s_{60}^{(2)} \approx 1.4$  for the two populations analyzed (years per 1,000 people of age 60 or more). Hence, standard errors estimated for the actual population sizes would be

$$SD_{60}^{(1)} = 1.5 / \sqrt{\frac{5930}{1000}} = 0.6 \text{ and } SD_{60}^{(2)} = 1.4 / \sqrt{\frac{15629}{1000}} = 0.3 \text{ (years).}$$

Assuming the independence of the estimates for the two populations, we may compute the standard error of the difference between the estimates of life expectancy:

$$SD_{60}^{(1)-(2)} = \sqrt{(SD_{60}^{(1)})^2 + (SD_{60}^{(2)})^2} = 0.7 \text{ (years)}.$$

Given the standard error and assuming normal distribution, the observed difference of  $26.1-25.5=0.6$  years yields  $p$ -value 40% (double-sided alternative), that is, *the difference may not be considered significant* at the 5% significance level. The two populations do not differ significantly with respect to life expectancy at age 60.

### Case 3. Minimal population size meeting the required level of estimation accuracy

Consider a situation where life expectancy at age 60 is estimated to be about 25 years, the proportion of the population aged 60 and more is 30%, and the age composition is stationary. Then suppose that the policymaker demands measurements of life expectancy at age 60 to be made at the regional level, with errors not exceeding 0.75 years at a 95% confidence level. What would the recommendation be about minimal population size for estimating the life expectancy at age 60 with the required accuracy? A difference of 0.75 years would not be statistically significant at the 95% confidence level at a standard error higher than  $\frac{0.75}{1.96} = 0.38$  years (assuming normal distribution, double-sided hypothesis). For a stationary population with  $e_{60} = 25$ , eq. (9) implies that

$s_{60} = 0.082 \cdot 25 - 0.001 \cdot 25^2 \approx 1.43$  years, that is, the critical threshold 0.38 of standard error may be reached at population size  $N_{60+} = \left(\frac{1.43}{0.38}\right)^2 = 14$  (thousand) at age 60 or

higher, that is, at total population size  $N = \frac{14}{0.3} \approx 46$  (thousand). Hence, estimation of life expectancy at 60 may be recommended for areas with at least 46,000 people.

## 6. General recommendations

We have shown that both the standard errors and the estimation bias become very high at a population size of around 5,000 or less. Additionally the distributions of standard errors deviate strongly from normality at such population sizes, which precludes building confidence limits and conducting other statistical analyses. Therefore estimating life expectancies for such populations must be discouraged.

Based on Appendix Table A1 and assuming that the standard error of the estimates of the life expectancy at birth is about one year or less, we may conclude that population exposure years should be about 15,000 people or more for a low-mortality population. To estimate life expectancy at 60 with a standard error of about 0.25 years, the population size should be about 100,000 or more for stationary populations, 65,000 for populations declining at 2%, and 250,000 for populations growing at 2% per annum. These rough estimates only outline how strict the requirements regarding population size could be to secure relatively accurate estimations.

We found the age composition of the population to be important for the accuracy of estimating life expectancy in small populations.

Precise assessments of standard errors and of minimal population size may vary considerably depending on actual population age composition and mortality schedules. Even the requirements for standard errors may vary from population to population, depending, for example, on observed spatial and social variation of mortality as well as on policy demands. In a country with high spatial diversity in life expectancy (e.g., Russia), even a low-precision estimate of life expectancy at the municipal level may reveal important regional differences, while for a country with more homogeneous regional mortality variation, like many western European countries, estimates must be conducted with higher precision, so that they reveal informative variations of mortality levels and not the random sample-size effects.

In most applications of standard errors, it is convenient to assume a normal distribution of the estimates. Our simulations indicate that such assumptions may safely be used starting from a population size of 50,000 people. For populations of 5,000 or less, such assumptions are not acceptable. In intermediate situations, normality assumptions may be used only as a rough approximation. More precise assessments, if necessary, may demand a detailed analysis and perhaps additional simulations tailored to the particular situation.

We do not find any advantages in using abridged life tables instead of unabridged ones even for a small population with many age groups containing no death observations. What is more, abridged life table calculations may lead to strong biases when the population age composition deviates from the stationary composition. Hence, it might well be advisable to use the unabridged life tables rather than the abridged ones when the population is not stationary.

Our simulation results show that procedures for the open age interval are crucial for the efficiency of life expectancy estimation. Although we found efficiency in our simple approach based on adjusting the open age interval in such a way that there is at least one death observed, more research on procedures for the open age interval could be important.

## **7. Acknowledgements**

We thank the editor and reviewers for helpful comments.

## References

- Ahn, S. and Fessler, A. (2003). *Standard Errors of Mean, Variance, and Standard Deviation Estimators*. Technical Report. Ann Arbor, MI, USA: EECS Department, University of Michigan: July 2003. <http://www.eecs.umich.edu/~fessler/papers/lists/files/tr/stderr.pdf>.
- Chiang, Ch.L. (1984). *The life table and its applications*. Malabar, Florida: Robert E. Krieger Publishing Company.
- Eayres, D. and Williams, E.S. (2004). Evaluation of methodologies for small area life expectancy estimation. *Journal of Epidemiology & Community Health* 2004(58): 243-249. doi:10.1136/jech.2003.009654.
- Fries, J.F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine* 303(3):130-135.
- Human Mortality Database* (2010). Database supported by University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). <http://www.mortality.org>; <http://www.humanmortality.de>.
- Keyfitz, N. (1977). *Applied mathematical demography*. John Wiley & Sons.
- Rao, J.N.K. (1973). *Linear statistical inference and its applications, 2nd Edition*. New York: Wiley.
- Silcocks, P.B.S., Jenner, D.A., and Reza, R. (2001). Life expectancy as a summary of mortality in a population: Statistical considerations and suitability for use by health authorities. *Journal of Epidemiology & Community Health* 2001(55): 38-43. doi:10.1136/jech.55.1.38.
- Toson, B., Baker, A., and the Office of National Statistics (2003). Life expectancy at birth: Methodological options for small populations. *National Statistics Methodological Series* 2003(33): 1-27. [http://www.statistics.gov.uk/downloads/theme\\_other/GSSMethodology\\_No\\_33.pdf](http://www.statistics.gov.uk/downloads/theme_other/GSSMethodology_No_33.pdf).
- Williams, E., Dinsdale, H., Eayres, D., and Tahzib, F. (2005). *Technical Report: Calculating Life Expectancy in small areas*. South East Public Health Observatory.

## Appendix: Supplementary tables

**Table A1: Standard for life expectancy estimates at different life expectancy, population growth rate, and size (see explanatory notes below the table)**

Range of life expectancy at birth, years	Population size 1,000 persons			Population size >= 5,000 persons		
	$SD_0$	$SD_{60}$	$s_{60}$	$s_0$	$s'_{60}$	$s_{60}$
<i>At population growth rate 0%:</i>						
35 or less	5.6 (0.5)	4.2 (0.3)	1.4 (0.1)	5.3 (0.5)	2.6 (0.2)	0.9 (0.1)
35-40	6.3 (0.2)	3.8 (0.2)	1.3 (0.1)	6.0 (0.2)	2.4 (0.1)	0.9 (0.0)
40-45	6.5 (0.2)	3.8 (0.2)	1.4 (0.1)	6.3 (0.2)	2.4 (0.1)	0.9 (0.1)
45-50	6.8 (0.2)	3.8 (0.2)	1.4 (0.1)	6.5 (0.2)	2.5 (0.1)	0.9 (0.1)
50-55	7.0 (0.2)	3.7 (0.2)	1.5 (0.1)	6.7 (0.2)	2.5 (0.1)	1.0 (0.1)
55-60	6.9 (0.2)	3.6 (0.2)	1.5 (0.1)	6.6 (0.2)	2.5 (0.1)	1.0 (0.1)
60-65	6.4 (0.4)	3.5 (0.2)	1.5 (0.1)	6.1 (0.4)	2.5 (0.1)	1.1 (0.1)
65-70	5.6 (0.4)	3.4 (0.2)	1.5 (0.1)	5.3 (0.5)	2.4 (0.1)	1.1 (0.1)
70-75	5.0 (0.3)	3.4 (0.2)	1.6 (0.1)	4.6 (0.4)	2.5 (0.1)	1.1 (0.1)
75-80	4.7 (0.2)	3.5 (0.3)	1.7 (0.1)	4.2 (0.2)	2.6 (0.1)	1.3 (0.1)
80 and more	4.4 (0.1)	3.4 (0.2)	1.8 (0.1)	3.9 (0.1)	2.6 (0.1)	1.4 (0.1)
<i>At population growth rate -1%:</i>						
35 or less	6.2 (0.7)	2.9 (0.3)	1.1 (0.1)	6.0 (0.6)	2.1 (0.1)	0.8 (0.1)
35-40	7.0 (0.3)	2.6 (0.1)	1.1 (0.0)	6.8 (0.3)	2.0 (0.1)	0.9 (0.0)
40-45	7.3 (0.3)	2.6 (0.1)	1.1 (0.1)	7.1 (0.3)	2.1 (0.1)	0.9 (0.0)
45-50	7.5 (0.3)	2.6 (0.1)	1.2 (0.1)	7.4 (0.3)	2.1 (0.1)	0.9 (0.1)
50-55	7.7 (0.2)	2.5 (0.1)	1.2 (0.1)	7.6 (0.2)	2.1 (0.1)	1.0 (0.1)
55-60	7.5 (0.3)	2.5 (0.1)	1.2 (0.1)	7.4 (0.3)	2.1 (0.1)	1.0 (0.1)
60-65	6.9 (0.5)	2.5 (0.1)	1.2 (0.1)	6.8 (0.5)	2.1 (0.1)	1.1 (0.1)
65-70	5.8 (0.6)	2.4 (0.1)	1.2 (0.1)	5.8 (0.6)	2.1 (0.1)	1.1 (0.0)
70-75	4.9 (0.6)	2.4 (0.1)	1.3 (0.1)	4.9 (0.6)	2.1 (0.1)	1.2 (0.1)
75-80	4.3 (0.3)	2.5 (0.1)	1.4 (0.1)	4.3 (0.3)	2.2 (0.1)	1.3 (0.1)
80 and more	3.9 (0.2)	2.4 (0.1)	1.5 (0.1)	4.0 (0.2)	2.2 (0.1)	1.4 (0.1)
<i>At population growth rate 1%:</i>						
35 or less	5.3 (0.4)	5.2 (0.3)	1.4 (0.2)	4.9 (0.4)	3.2 (0.3)	0.9 (0.1)
35-40	5.9 (0.2)	5.1 (0.2)	1.5 (0.1)	5.5 (0.2)	3.0 (0.2)	0.9 (0.1)
40-45	6.2 (0.2)	5.2 (0.2)	1.5 (0.1)	5.7 (0.2)	3.0 (0.2)	0.9 (0.1)
45-50	6.6 (0.2)	5.4 (0.3)	1.7 (0.1)	6.0 (0.1)	3.0 (0.2)	1.0 (0.1)
50-55	6.9 (0.2)	5.5 (0.3)	1.8 (0.1)	6.2 (0.1)	3.0 (0.2)	1.0 (0.1)
55-60	6.9 (0.2)	5.5 (0.3)	1.9 (0.1)	6.1 (0.2)	3.0 (0.2)	1.0 (0.1)
60-65	6.7 (0.3)	5.5 (0.3)	2.0 (0.1)	5.7 (0.3)	3.0 (0.1)	1.1 (0.1)
65-70	6.3 (0.3)	5.4 (0.3)	2.0 (0.1)	5.1 (0.3)	2.9 (0.1)	1.1 (0.1)
70-75	6.2 (0.3)	5.6 (0.3)	2.2 (0.1)	4.6 (0.3)	3.0 (0.1)	1.2 (0.1)
75-80	6.3 (0.3)	5.9 (0.4)	2.4 (0.2)	4.3 (0.1)	3.1 (0.1)	1.3 (0.1)
80 and more	6.4 (0.3)	6.0 (0.3)	2.7 (0.2)	4.1 (0.1)	3.1 (0.1)	1.4 (0.1)

**Table A1: (Continued)**

Range of life expectancy at birth, years	Population size 1,000 persons			Population size >= 5,000 persons		
	$SD_0$	$SD_{60}$	$s_{60}$	$s_0$	$s'_{60}$	$s_{60}$
<i>At population growth rate -2%:</i>						
35 or less	7.3 (0.9)	2.1 (0.1)	1.0 (0.1)	7.0 (0.8)	1.8 (0.1)	0.9 (0.1)
35-40	8.3 (0.4)	2.0 (0.1)	1.0 (0.0)	8.0 (0.4)	1.8 (0.1)	0.9 (0.0)
40-45	8.6 (0.4)	2.0 (0.1)	1.0 (0.1)	8.3 (0.4)	1.8 (0.1)	0.9 (0.0)
45-50	8.9 (0.4)	2.0 (0.1)	1.1 (0.0)	8.7 (0.4)	1.8 (0.1)	1.0 (0.1)
50-55	9.1 (0.3)	2.0 (0.1)	1.1 (0.1)	9.0 (0.3)	1.8 (0.1)	1.0 (0.1)
55-60	8.9 (0.5)	2.0 (0.1)	1.1 (0.1)	8.8 (0.4)	1.9 (0.1)	1.1 (0.1)
60-65	7.9 (0.7)	2.0 (0.1)	1.1 (0.1)	8.0 (0.6)	1.9 (0.1)	1.1 (0.1)
65-70	6.7 (0.9)	2.0 (0.1)	1.1 (0.0)	6.8 (0.8)	1.8 (0.1)	1.1 (0.1)
70-75	5.4 (0.8)	2.0 (0.1)	1.2 (0.1)	5.6 (0.8)	1.9 (0.1)	1.2 (0.1)
75-80	4.6 (0.4)	2.1 (0.1)	1.3 (0.1)	4.8 (0.4)	2.0 (0.1)	1.3 (0.1)
80 and more	4.1 (0.2)	2.1 (0.1)	1.4 (0.1)	4.3 (0.2)	2.0 (0.1)	1.4 (0.1)
<i>At population growth rate 2%:</i>						
35 or less	5.0 (0.4)	5.0 (0.5)	1.1 (0.2)	4.7 (0.3)	4.3 (0.7)	0.9 (0.1)
35-40	5.6 (0.2)	5.3 (0.3)	1.3 (0.1)	5.2 (0.1)	3.9 (0.5)	0.9 (0.1)
40-45	6.0 (0.2)	5.4 (0.4)	1.3 (0.1)	5.5 (0.2)	3.9 (0.4)	0.9 (0.1)
45-50	6.4 (0.2)	5.9 (0.3)	1.5 (0.1)	5.8 (0.1)	3.9 (0.4)	1.0 (0.1)
50-55	6.9 (0.2)	6.4 (0.3)	1.8 (0.1)	6.0 (0.2)	3.8 (0.4)	1.0 (0.1)
55-60	7.2 (0.3)	6.7 (0.4)	1.9 (0.1)	6.0 (0.2)	3.8 (0.4)	1.1 (0.1)
60-65	7.2 (0.4)	6.8 (0.4)	2.0 (0.2)	5.7 (0.3)	3.8 (0.3)	1.1 (0.1)
65-70	7.1 (0.4)	6.8 (0.3)	2.0 (0.2)	5.3 (0.3)	3.7 (0.3)	1.1 (0.1)
70-75	7.5 (0.3)	7.4 (0.4)	2.4 (0.2)	5.0 (0.2)	3.8 (0.3)	1.2 (0.1)
75-80	8.2 (0.4)	8.2 (0.5)	2.8 (0.2)	4.8 (0.2)	3.9 (0.3)	1.3 (0.1)
80 and more	8.8 (0.5)	8.8 (0.5)	3.2 (0.2)	4.7 (0.2)	3.9 (0.2)	1.4 (0.1)

Notes:  $SD_0$ ,  $SD_{60}$  - standard error of the estimated life expectancy at birth and at age 60, respectively;  $s_{60} = SD_{60} \sqrt{N_{60} / 1000}$  - standard error of the estimated life expectancy at age 60 rescaled to population size 1,000 at age 60 and above;  $s_0 = SD_0 \sqrt{N / 1000}$ ,  $s'_{60} = SD_{60} \sqrt{N / 1000}$  - standard error of the estimated life expectancy at birth and at age 60, respectively, rescaled to total population size 1,000. Standard errors are obtained by averaging sample standard deviations (rescaled for populations above 1,000 people) over all populations with life expectancies in a given range. Numbers in parentheses represent the standard error of indicators over the entire set of simulated populations. Results for Russian life tables are excluded while averaging because of considerably different patterns.



**Table A2: Percentiles of distribution of estimates for life expectancy at birth, derived from simulated distribution and from corresponding normal distribution (simulations are based on stable populations and mortality schedules corresponding to the life table of the Japanese female population in 2007)**

Population size	Growth rate		Percentile:					
			2.5%	5.0%	50.0%	95.0%	97.5%	
5000	-2%	Normal distribution	82.42	83.01	86.09	89.16	89.75	
		Actual distribution	81.66	82.61	86.33	88.67	89.08	
	-1%	Normal distribution	82.69	83.24	86.10	88.96	89.50	
		Actual distribution	82.31	83.03	86.23	88.74	89.19	
	0%	Normal distribution	82.70	83.26	86.15	89.04	89.60	
		Actual distribution	82.59	83.17	86.21	88.94	89.49	
	1%	Normal distribution	82.49	83.10	86.31	89.51	90.12	
		Actual distribution	82.46	83.14	86.30	89.44	90.03	
	2%	Normal distribution	81.80	82.55	86.46	90.37	91.12	
		Actual distribution	82.10	82.81	86.38	90.29	91.22	
	10000	-2%	Normal distribution	83.45	83.87	86.03	88.18	88.60
			Actual distribution	83.15	83.70	86.15	87.97	88.25
-1%		Normal distribution	83.65	84.04	86.05	88.06	88.44	
		Actual distribution	83.51	83.96	86.11	87.92	88.25	
0%		Normal distribution	83.69	84.07	86.08	88.09	88.47	
		Actual distribution	83.58	84.00	86.11	88.04	88.39	
1%		Normal distribution	83.49	83.91	86.14	88.36	88.79	
		Actual distribution	83.40	83.87	86.15	88.34	88.74	
2%		Normal distribution	83.14	83.64	86.23	88.83	89.32	
		Actual distribution	83.24	83.71	86.22	88.88	89.36	
50000		-2%	Normal distribution	84.82	85.01	85.99	86.98	87.17
			Actual distribution	84.77	84.97	86.02	86.93	87.11
	-1%	Normal distribution	84.91	85.09	86.01	86.93	87.10	
		Actual distribution	84.87	85.06	86.02	86.89	87.07	
	0%	Normal distribution	84.94	85.11	86.01	86.91	87.08	
		Actual distribution	84.92	85.11	86.02	86.90	87.06	
	1%	Normal distribution	84.83	85.02	86.00	86.99	87.17	
		Actual distribution	84.83	85.02	86.01	86.99	87.18	
	2%	Normal distribution	84.69	84.90	86.03	87.16	87.38	
		Actual distribution	84.70	84.90	86.02	87.16	87.40	
	100000	-2%	Normal distribution	85.17	85.30	85.99	86.68	86.82
			Actual distribution	85.13	85.28	86.01	86.66	86.77
-1%		Normal distribution	85.21	85.33	85.99	86.64	86.76	
		Actual distribution	85.20	85.32	86.00	86.62	86.74	
0%		Normal distribution	85.23	85.35	85.99	86.63	86.75	
		Actual distribution	85.21	85.34	85.99	86.62	86.73	
1%		Normal distribution	85.17	85.30	85.99	86.68	86.82	
		Actual distribution	85.16	85.30	85.99	86.69	86.82	
2%		Normal distribution	85.08	85.23	86.01	86.80	86.95	
		Actual distribution	85.09	85.23	86.01	86.80	86.97	

**Table A3: Percentiles of distribution of estimates for life expectancy at age 60, derived from simulated distribution and from corresponding normal distribution (simulations are based on stable populations and mortality schedules corresponding to the life table of the Japanese female population in 2007)**

Population size	Growth rate		Percentile:					
			2.5%	5.0%	50.0%	95.0%	97.5%	
5000	-2%	Normal distribution	26.23	26.53	28.12	29.71	30.02	
		Actual distribution	26.24	26.53	28.14	29.70	30.00	
	-1%	Normal distribution	26.05	26.39	28.15	29.92	30.26	
		Actual distribution	26.08	26.41	28.14	29.92	30.28	
	0%	Normal distribution	25.77	26.17	28.23	30.29	30.68	
		Actual distribution	25.85	26.21	28.20	30.32	30.76	
	1%	Normal distribution	25.29	25.78	28.35	30.92	31.41	
		Actual distribution	25.42	25.88	28.31	30.93	31.45	
	2%	Normal distribution	24.45	25.11	28.55	31.99	32.65	
		Actual distribution	24.89	25.45	28.41	31.88	32.87	
	10000	-2%	Normal distribution	26.76	26.97	28.08	29.20	29.41
			Actual distribution	26.75	26.98	28.08	29.20	29.41
-1%		Normal distribution	26.64	26.87	28.10	29.33	29.57	
		Actual distribution	26.62	26.85	28.10	29.33	29.55	
0%		Normal distribution	26.40	26.68	28.13	29.57	29.85	
		Actual distribution	26.42	26.68	28.12	29.56	29.86	
1%		Normal distribution	26.12	26.45	28.20	29.94	30.27	
		Actual distribution	26.16	26.48	28.18	29.96	30.31	
2%		Normal distribution	25.67	26.09	28.31	30.52	30.95	
		Actual distribution	25.80	26.19	28.26	30.56	31.06	
50000		-2%	Normal distribution	27.46	27.56	28.06	28.56	28.66
			Actual distribution	27.46	27.56	28.06	28.57	28.67
	-1%	Normal distribution	27.40	27.51	28.07	28.62	28.73	
		Actual distribution	27.41	27.51	28.06	28.62	28.73	
	0%	Normal distribution	27.33	27.45	28.08	28.71	28.83	
		Actual distribution	27.31	27.45	28.08	28.71	28.84	
	1%	Normal distribution	27.16	27.30	28.07	28.84	28.98	
		Actual distribution	27.17	27.30	28.07	28.84	28.99	
	2%	Normal distribution	26.96	27.14	28.10	29.05	29.24	
		Actual distribution	26.97	27.15	28.09	29.07	29.26	
	100000	-2%	Normal distribution	27.64	27.71	28.06	28.41	28.48
			Actual distribution	27.63	27.70	28.06	28.41	28.48
-1%		Normal distribution	27.59	27.66	28.06	28.45	28.52	
		Actual distribution	27.58	27.66	28.06	28.44	28.51	
0%		Normal distribution	27.53	27.61	28.06	28.51	28.59	
		Actual distribution	27.53	27.61	28.06	28.51	28.59	
1%		Normal distribution	27.42	27.52	28.06	28.60	28.70	
		Actual distribution	27.43	27.52	28.06	28.61	28.70	
2%		Normal distribution	27.29	27.41	28.08	28.75	28.88	
		Actual distribution	27.30	27.43	28.08	28.76	28.88	