On the Accuracy of Life Expectancy

Nan Li, United Nations[1]
Shripad Tuljapurkar, Stanford University

Publishing annual life expectancy by sex to two decimals has become almost customary for countries with reliable death registration, and perhaps a target for other nations. The basis of this standard requires investigation. Following common practice in statistics, we define accuracy as the 95% confidence interval, and provide a simple method to compute the accuracy of calculations of life expectancy. We show that the accuracy of life expectancy is mainly determined by the level of mortality and the size of population. We indicate that, even if death registration and population count were perfect, the accuracy of life expectancy would not reach a year for 30% of all countries, 0.1 years for 63% of all countries, and 0.01 years for any country, even China or India.

Life expectancy at birth (hereafter life expectancy, *LE*) is increasingly used as a broad measure of mortality, and applied to issues such as evaluating the development levels of countries (e.g., UNDP, 2010). Given this trend, it is practically important to ask how accurately the LE can be computed.

International organizations estimate and publish the *LE*, for countries with or without death registration, in a variety of formats. The United Nations Population Division (UNPD, e.g., United Nations, 2009) publishes the *LE* by sex to one decimal, for 5-year periods and countries with 100 thousand or more populations. The United Nations Development Programme (e.g., UNDP, 2010) computes Human Development Index using the *LE* estimated by UNPD, but for two sexes together and to one decimal. The World Health Organization (e.g., WHO, 2005) estimates *LE* by sex to integer years, but for single-year periods. On the other hand, almost every country, or a sub-area of a country that computes life tables, publishes annual *LE* by sex to two decimals (e.g., Sweden, 2011; Japan, 2011). In general, international organizations view their estimates of *LE* as less accurate than that of the national statistical agencies. We have not been able to trace the origin of these standards nor their rationale. One might make the common sense argument that it is reasonable to expect that countries with good vital statistics can register all their deaths accurately to one day. Then we might expect the *LE* of these countries to be accurate to the second decimal, which corresponds to 0.01 years or 3.7 days.

But this argument doesn't answer to the question: can the *LE* for a year be computed accurately by sex to two decimals for any country? To answer this question, "accurate" needs to be defined. Following common practice in statistics, we define accuracy as the 95% confidence interval. In other words, an accuracy of 0.01 years implies that there is a 95% chance for the true *LE* to be in the range of 0.005 years around the computed *LE*. Under this definition, as we will show in this paper, the answer to the above question is negative: indeed *LE* cannot be accurately computed to 0.01 years annually by sex for any country, even China or India.

[1] The views expressed in this paper are those of the author's and do not necessarily reflect those of the United Nations.

Among the methods of estimating the accuracy of *LE*, Chiang (1984) provided the following approach. Assuming that the age-specific probabilities of survival are measured without bias and that deaths are binomially distributed within an age group, Chiang first derived the formula to compute the standard error of survival probability of an age group, which depends on the observed death rate and the number of death in this age group. Describing the change of *LE* by survival probabilities, Chiang obtained the formulas of computing the standard error of *LE*, which depends on the age-specific death rates and deaths. Using these formulas to the data of US females in 1975, Chiang showed that the standard error of *LE* is about 0.016, leading to 0.06 years accuracy in terms of normal distribution. Chiang's work indicated that when the death rates are assumed unbiased, population size is the main factor that determines the accuracy of *LE*. Moreover, even for the population of the US in 1975, the third largest one after China and India then, the LE could not be computed accurately to 0.01 years. The logic behind Chiang's study can be described by the large number law: the chance of seeing the face from throwing a coin is 0.5 in theory, but in practice one can make it close to 0.5 only when number of throws is large.

Assuming that the distribution of deaths within an age group is of Poisson rather than binomial, Silcocks et al (2001) provided another set of formulas to estimate the accuracy of *LE*. WHO (2005) applied a method similar to that of Silcocks to estimate the accuracy of *LE* for countries with complete death registration, and used a simulation method that is based on the uncertainty of model life table parameters to estimate the accuracy of *LE* for other countries. These methods, however, are complex (Silcocks et al, 2001). Nonetheless, Eayres and Willions (2004) applied the methods of Chiang and Silcocks to hypothetical small populations with age structure and death rates of English men in 1998-2000, and obtained more impressive results. They showed, for instance, the standard error of *LE* for a hypothetical population of 50 thousands is about 0.6 years, implying that the accuracy is more than 2 years. Their study suggested that, if the *LE* is required to be accurate to 2 years, then it would not be reachable for populations less than 50 thousand for one sex, or for countries with less than 100 thousand populations. These conclusions are informative, but they are based on either simulation for small populations or death rates and age structure of a chosen country. Can more general conclusions be obtained by some method easier to understand, and simpler to use? The answer is yes.

A one-birth model

We start from describing the life circle of one birth rather than the deaths in one age group. We denote the years this person lives by a random variable Y_1, where subscript 1 stands for the first person for the reason to be seen soon. Let the probability of surviving from birth to age *x* be ${}_x p_0$ and the probability of dying in the age group with starting age *x* be q_{x+n}, where n is the length of the age group, then the probability of surviving from birth to age x and dying in age group x, namely $P\{Y_1 = \bar{x}\}$, is

$$P\{Y_1 = \bar{x}\} = \delta_x = {}_x p_0 q_{x+n}, \qquad (1)$$

where $\bar{x}$ is the average age of death in age group *x*, which is described in the appendix. When the $\bar{x}$ is chosen properly, how to cut the age groups should not matter, and hereafter we focus on the age groups of abridged life tables. In (1), the δ_x is the density function of a probability distribution of death, which is obtained from observed death rates, and can be called the age pattern of death. It is worth noting that the basis of our method is this probability distribution over all age groups. On the other hand, the method of (Chiang, 1984) or (Silcoks, 2001) is based on the probability distribution of deaths within each group.

Given the δ_x, the life expectancy of the person in question is

$$LE = E(Y_1) = \sum_{x=0}^{\omega} \bar{x} P\{Y_1 = \bar{x}\} = \sum_{x=0}^{\varpi} \bar{x}\delta_x, \qquad (2)$$

where ω is the starting age of the oldest age group, which is described in the appendix. Subsequently, the standard deviation of Y_1, or the standard error of estimating the LE of the one-birth case, namely S, is

$$S = \sqrt{V(Y_1)} = \sqrt{\sum_{x=0}^{\omega} (\bar{x} - LE)^2 P\{Y_1 = \bar{x}\}} = \sqrt{\sum_{x=0}^{\omega} (\bar{x} - LE)^2 \delta_x} \quad (3)$$

One may note that zero death in some age group does not cause problem here, but yields infinite standard deviation in previous methods.

As is shown by Edwards and Tuljapurkar (2005), the values of S observed from seven developed countries in the last 50 years ranged from 14 to 20 years. Thus, estimating the *LE* from one person would result in unacceptable accuracy, and we move forward to a cohort model.

A cohort model

Here we consider *B* persons born in one year and subject to the same mortality at all the ages, and denote the years that the ith person lives by a random variable Y_i. Then, the average years that the *B* persons live, namely the random variable $\bar{Y}$, is:

$$\bar{Y} = \frac{\sum_{i=1}^{B} Y_i}{B}. \qquad (4)$$

Now the situation is entirely different. First, all Y_i obey the same distribution and they are independent each other. Second, according to the central limit theorem, the distribution of $\bar{Y}$ will be close to a normal distribution no matter what is the empirical distribution of Y_i. And third, the number of Y_i refers to the annual births of a country that is sufficient larger than 25 or 30, bigger than which the distribution of $\bar{Y}$ will be satisfactorily close to a normal distribution (e.g., see Agresti and Finlay, 1997, p104). These features make the confidence interval referred by standard error valid. As a comparison, one may recall that in the previous methods the binomial or Poisson distributions differ significantly over age, and the number of these distributions is usually around 20 and can hardly be larger.

We now turn to compute the expectation and standard deviation of $\bar{Y}$. The expectation of $\bar{Y}$ is still (2), and the standard error of estimating LE is:

$$\sqrt{V(\bar{Y})} = \sqrt{V(\frac{\sum_{i=1}^{B} Y_i}{B})} = \frac{S}{\sqrt{B}}. \qquad (5)$$

This is the standard error of estimating the LE for the *B* births, which is equivalent to observing the average age of deaths following the life circles of the B births. But in practice one can rarely follow the life circle of a cohort, and we therefore turn to a stationary population model.

A stationary population model

A deterministic stationary population will be reached by any initial population, if for any year the number of births is *B*, the probabilities of death are q_x, and there is no migration. The basis of the deterministic stationary population is the large number law, with which the 0.9 probability of surviving to age 5, for example, is interpreted as 90% births will survive to age 5. Without applying the large number law, the deterministic stationary population becomes probabilistic, in which deaths occur independently between cohorts. This stationary population has the age structures of death and population that are identical to the above cohort model. Thus, observing the average age of deaths following the above cohort model is equivalent to doing so among the corresponding stationary population. More specifically, these deaths are to be observed among the stationary population with $P = B \cdot LE$ persons in one year[2]. Denote the accuracy of the LE by Ac, (5) leads to

$$Ac/2 = 1.96S/\sqrt{B} = 1.96S\sqrt{LE}/\sqrt{P}. \qquad (6)$$

Now the question is whether the real population size can be used as *P*. From a statistic point of view, *LE* is the mean of the average age of deaths among a stationary population ($LE = E(\bar{Y})$), and is independent with the size of the stationary population. Thus, in constructing deterministic life tables, the number of births is often chosen as 100 thousands for no clear reason, because the size of the stationary population does not matter when only a non-biased estimate of *LE* is interested. When the error of estimating *LE* is in question, however, the size of the stationary population, namely *P*, matters, and how to choose it is a question. We do not suggest simply use the real population size as *P*, because *LE* is the mean of the average age of deaths among a stationary population, not among a real population. We suggest, instead, choose the *P* as the size of the stationary population that is the closest to the real population. Let $p_o(x)$ be the observed population in age group *x*, and L_x be the person-years in age group *x* of the stationary population, which are the commonly used life table function starting from 100 thousand births. Our suggestion leads to minimize $\sum_{x=0}^{\omega}[p_o(x) - P\frac{L_x}{\sum_{x=0}^{\omega} L_x}]^2$, which results in

[2] In a probabilistic stationary population, the total person-years is a random variable and $P = B \cdot LE$ is its expectation.

$$P = \frac{\sum_{x=0}^{\omega} p_o(x) \frac{L_x}{\sum_{x=0}^{\omega} L_x}}{\sum_{x=0}^{\omega} (\frac{L_x}{\sum_{x=0}^{\omega} L_x})^2} = \frac{\sum_{x=0}^{\omega} L_x \sum_{x=0}^{\omega} p_o(x) L_x}{\sum_{x=0}^{\omega} L_x^2}. \qquad (7)$$

If the real population is stationary, (7) leads to the real population size. Accordingly, probabilistic life tables can be constructed, of which the size of a cohort at birth, namely l_o, should be given by the equivalent stationary population's births *B(=P/LE)*, and it should not be arbitrarily taken as 100 thousand for all the countries at all the times. Starting from the country-time-specific number of birth, $l_o = B$, each person of the cohort will survive or die randomly according to the probability of death at each age, q_x. Consequently, at each age, all life table variables have probability distributions, of which the central tendencies compose that of the equivalent stationary population or a deterministic life table. Therefore, the nature of life tables is probabilistic, and deterministic life tables are approximations for large populations.

It is clear that when a cohort follows the age-specific death rates of a real population to survive, it survives over age randomly according to the age-specific variances determined by the underlying stationary population, which is often not the real population. In other words, a cohort cannot survive over age randomly according to the age-specific variances that are determined by populations other than the cohort itself. On the other hand, when the age-specific variances generated by an observed population are used to compute the variance of *LE*, as is implemented by Chiang's method, the result is not the variance of the age of deaths in a cohort's survival process, but something else that is hard to interpret. More specifically, Chiang's formulas of computing the variance of *LE* can be viewed as a weighted average of age-specific variances, in which the weights are computed using the stationary population while the age-specific variances are calculated by observed deaths. Thus, there is a logic inconsistence in Chiang's formulas, because a stationary population and observed deaths cannot often exist together. This inconsistence can be eliminated by using Chiang's formula on the deaths of a corresponding stationary population rather than an observed population. By doing so, the result of Chiang's formula will be identical to ours, as is shown in the appendix. Despite the logic inconsistence, there are facts that would make the results of Chiang practically close to ours. First, the variance of life expectancy is determined by the age pattern of death rather than population. Second, the age patterns of death are more similar between populations than are the age patterns of population. Thus, the results Chiang should be similar to ours, as will be shown by examples.

The situation is simpler when ask what is the stationary population size in order to reach a certain *Ac*. It is

$$P = [1.96S\sqrt{LE}/(Ac/2)]^2. \qquad (8)$$

Note that a stationary population may exist in any length of period, the above accuracy can be reached by *P* persons in one year, or by *P/t* persons in *t* years, where *t* can be any positive number. Apparently, to reach the same accuracy, the persons in one year could be reduced to 1/5 if the period is extended to 5 years. On the other hand, however, the precision of referring time is reduced similarly. Hereafter, the persons in one year are used in this paper, and are simply called the population size.

Given q_x, *S* is obtained directly from (3), and then the *Ac* can be computed by (6) and (7), or the *P* for a required *Ac* can be calculated by (8). The basic indicator is *S*, which measures the uncertainty caused by finite population size when the death registration and population count were perfect. When there are errors in death registration and population count, the standard error in estimating *LE* should be larger than *S*, assuming that the that these errors are independent from that caused by finite population size. The above direct calculation, however, can hardly answer the question about how the *Ac* changes with *LE* in a more general sense. We therefore propose the indirect calculation below.

The indirect calculation

To maximize generality, we choose the West family of the Coale-Demeny (1966) model life tables (CDW) to represent the most common age pattern of mortality[3]. The values of S, computed at selected levels of *LE*, are shown in Table 1.

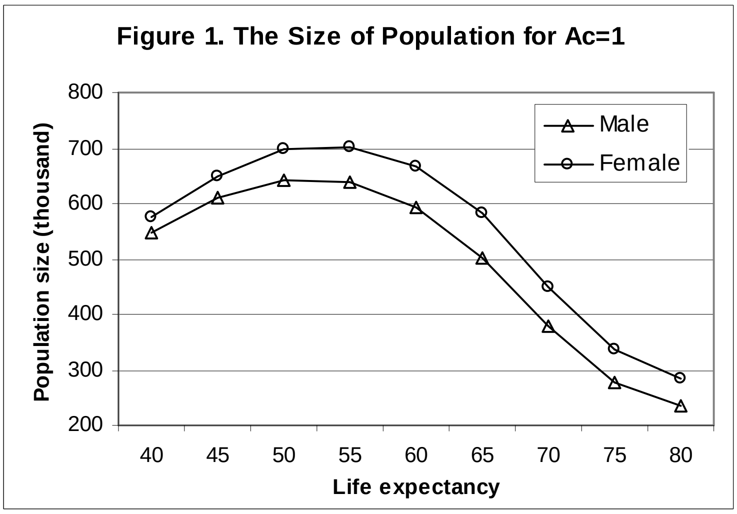
Table 1. The values of S computed using CDW

LE	40	45	50	55	60	65	70	75	80
Male S	29.9	29.7	28.9	27.5	25.4	22.4	18.8	15.5	13.8
Female S	30.6	30.7	30.1	28.8	26.9	24.2	20.5	17.1	15.2

It can be seen that the S declines with *LE*. The reason for this trend is understandable. There are two peaks in the distribution of death by age: one is at infant and another at old ages. With the increase of *LE*, the one at infant age drops and the other at old age raises, resulting in a decline of standard deviation. It is also apparent that the S of females are higher than that of males at the same *LE*, and the gap increases with *LE*. These patterns need further analysis. But, it is interesting to note that, the gap obscures or disappears when comparing the male S at a certain *LE* with the female S at the *LE* that is 5 years higher.

Using the values of S in table 1 and (6)-(7) or (8), the *Ac*, or the *P* for required *Ac* can be computed analytically according to the levels of *LE* (with some interpolation when the *LE* does not end with 0 or 5), which we call the indirect calculation. Another application of the indirect calculation is to investigate how the *P* for a required *Ac* changes with *LE*, of which we show the results for *Ac=1* in Figure 1.

[3] Other model life tables could also be utilised when there is reason to do so.



It can be seen that the P for $Ac=1$ first rises when LE is lower than 50, and then drops when LE is higher than 55. This is because, as is shown in (8), the P for required AC is a function of S^2LE, in which the S declines with LE, but together they show a non-monotonous trend.

For countries without direct measure of q_x, indirect estimate of LE are often obtained by using model life tables on surveyed child and/or adult death rates. Assuming that the errors caused by indirectly estimating LE are independent from the errors caused by finite population size, the values of S obtained by the indirect calculation could be viewed as the lower bounds for countries without direct measure of q_x.

Table 1 and (8) could provide approximate but informative conclusions, by taking the LE as 70 years for males and 75 years for females, roughly the median levels[4] of the all the countries in 2005-2010 (United Nations, 2011). In order for the LE to be accurate to 1 year, the population size should be larger than 0.38 $(=[1.96*18.8\sqrt{70}/(1/2)]^2)$ million for males and 0.34 million for females, or the total population should be 0.76 millions assuming the numbers of male and female are equal. In 2010, about 30% of all the countries[5] in the world had a size of population less than 0.76 million. Thus, for these countries, even if their death registrations and population counts were perfect, the computed LE would not be accurate to one year, and the 95% confidence intervals would cover different values of the last integer of the computed LE. Further, requiring the LE to be accurate to 0.1 years will raise the population size to 76 million, fewer than which and more then 0.76 million there were 63% countries in 2010. For these countries, the 95% confidence intervals would cover different values of the first decimal of the computed LE. Furthermore, requiring LE to be accurate to 0.01 years would need the size of population to reach 7.6 billion, which was bigger than the population size of the world in 2010. Thus, even for the left 7% countries with more than 76 million populations, the LE is accurate to 0.1 but not 0.01 years.

[4] Taking from the 196 countries or areas with 100 thousand or more populations.
[5] Including the 34 countries or areas with less than 100 thousand populations, and assuming the distribution of the LE among these countries is similar to that of the others.

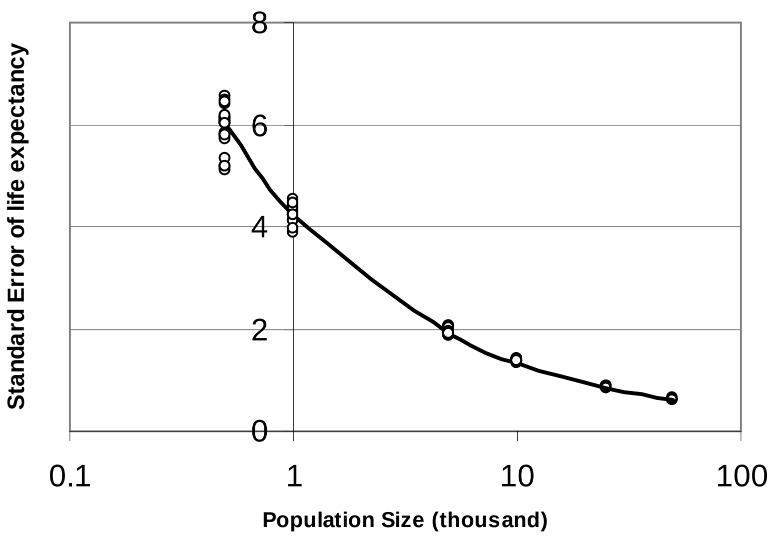
For about half of the countries, the *LE* are higher than the median, the required population size for the accuracy to reach 1 year, for example, is therefore smaller than 0.76 million, and hence the effect is to reduce the "30%" mentioned above. But there are other half countries with *LE* lower than the median, which bring an opposite effect. Thus, the above conclusions are inexact because the levels of *LE* differ among countries, but they are approximate because the effects from different *LE* should cancel each other. We should also mention that these conclusions are for annual *LE*. For the LE in 5-year periods, the required population size could be reduced to 1/5.

Although the indirect calculation is approximate in terms of using model rather than real life tables, it is useful for countries without direct measure of q_x. According to the WHO (2007), two-thirds of annual deaths are not registered. For countries without reliable death registration, the q_x could only be computed at some census years.

Examples and discussion

How does the indirect calculation perform? We compare its results with the previous studies. In the study of Eayres and Willions (2004), the standard errors of *LE* are computed using the q_x of English men in 1998-2000 (of which the *LE* is 75.42) and 6 hypothetical population sizes, which are 0.5, 1, 5, 10, 25, 50 thousands. For each population size, the age structure is taken from the English men in 1998-2000, and 30 standard errors of *LE* are computed, as are shown by the cycles in Figure 2. These 30 standard errors differed by using the methods of Chiang or Silcock, numerical computing or random simulating, the last age group starting at age 85 or 95, and 5-year or 10-year age group. For comparison, we use the indirect calculation with $S = 17.1$, which is chosen from table 1 at *LE=75*, the closest integer to the *LE* of English men in 1998-2000.

Figure 2. The Standard Error of Life Expectancy of Males: Curve(Indirect Calculation), Circles(Eayres and Willions)

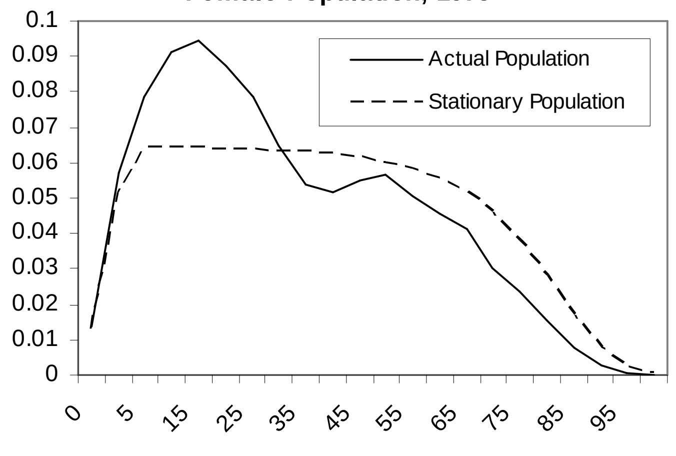


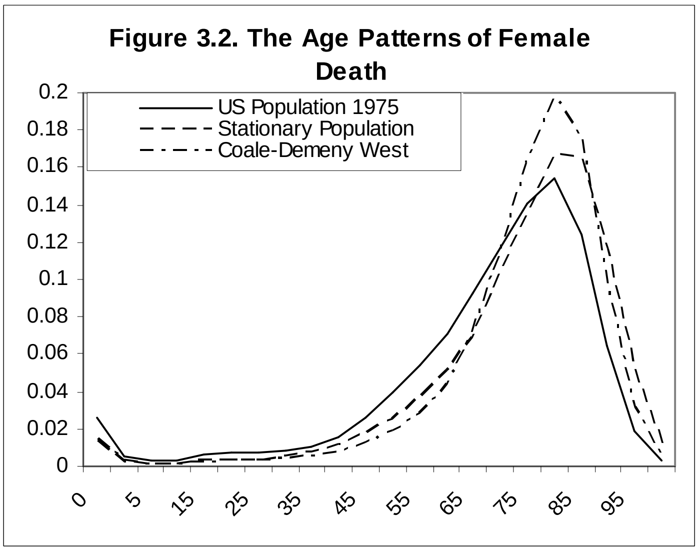
Eayres and Willions showed, in Figure 2, that the standard error of *LE* declines nonlinearly with *P*. The indirect calculation explained that this function is proportional to $1/\sqrt{P}$.

Now let us imagine a male population with the size of 50 thousand, and assume that perfect registration and calculation lead to *LE=75.42*, where the "2" is a result of rounding up following the customary standard. Is the second decimal accurate? And if not, what should be done? To answer these questions, we first find $S \approx 17$ according to Table 1, and then $Ac/2 = 1.28$ according to (6). We recommend rounding up the *Ac*/2 to its largest non-zero digit which is 1, and correspondingly the 75.42 to 75, then the *LE* is expressed as 75±1, reflecting the fact that the last integer is inexact. This example showed a case that the *LE* is inaccurate to 1 year, and the decimals could be misleading.

We now turn to an application of the direct calculation. Using the q_x and the female populations by age of the US in 1975, (7) produces the *P* as 114 millions, which is about 3.6% more than the actual population size. Further, (6) yields Ac/2=0.0289, and the indirect calculation that requires only the real population size and *LE* gives Ac/2=0.0271. Compared to that of Ac/2=0.0305 by Chiang's method, the direct calculation makes a difference about 0.0016 years. This difference is caused mainly by the age structures of the population, because the direct calculation counts the ages of death among a stationary population, while Chiang's method does so among an observed population. A comparison of the two age structures is shown in Figure 3.1. It is apparent that the baby boom and bust caused a significant difference between the two age structures. Compared to the difference between the age structures of population, the 0.0016 years difference between the values of Ac is small.

Figure 3.1. The Age Patterns of the US Female Population, 1975





How can the results between the methods of ours and that of Chiang be so close? We discuss some plausible reasons below. First, because the age pattern of death is determined not only by the age structure of population, but also regulated by the age pattern of mortality, the difference between the age patterns of death should be smaller than that between the populations, as can be seen by comparing in Figures 3.1 and 3.2. And because the CDW represented the most common age pattern of mortality, its age pattern of death should also be similar to the actual or the stationary one, as is also shown in Figure 3.2. Second, the variance of estimating *LE* is determined directly by the age pattern of death instead of population. Thus, the closeness between our results and that of Chiang is comprehensive.

Finally, for US females in 1975, about 111 millions then, the *LE* cannot be to 0.01 years, and it should be displayed as LE=76.65±0.03 years, an example that the first decimal is accurate but the second is not. The accurate first decimal suggests a concise format: the *LE* can be expressed as 76.7 years, with 95% sure that the last digit is exact. This conclusion was produced by a complex method using data on age-specific mortality and population, and it could be provided now by a simple method using data merely on the *LE* and total population. Furthermore, for countries with less population or higher levels of mortality, the *LE* can only be computed less accurately. We therefore suggest evaluate the accuracy when computing or estimating a *LE*.

Appendix

(1) The average age of deaths in an age group

As Eayres and Willions (2004) indicated, the average age of deaths in age group x, namely $\bar{x}$, is determined by assuming that the deaths are evenly distributed in one age group in Chiang's method, or by assuming that the death rate is constant in one age group in Silcock's method. It is more convenient to rewrite $\bar{x} = x + a_x$, where a_x is a commonly used life-table function, namely the mean number of person years lived by the deaths in age group x. In fact, Chiang's assumption leads to a_x being at the middle of the age group with length n, and Silcock's at the first half of the age group. Although these assumptions make little difference in computing

LE, they are too simple to be consistent with the qualitative features that a_x should follow. These features are described, for instance, by Preston, Heuveline and Guillot (2001, p44), as below.

First, a_x should be smaller than $\frac{n}{2}$ at infant and child ages. This is because, when the m_x (the death rate at age group x) is small and declines with age, more deaths should occur in the first half of the age interval. It is also clear that a_x should be larger than $\frac{n}{2}$ at adolescent and adult ages. This is because when m_x increases with age, more deaths should appear in the second half of the age interval. At old ages, a_x should be smaller than $\frac{n}{2}$, because m_x is high so that survivors decline quickly with age.

Among the formulas of estimating a_x, we recommend use the Greville (1943) formula as below:

$$a_x = \frac{n}{2} - \frac{n^2}{12}[m_x - \frac{Log(m_{x+n} / m_{x-n})}{2n}].$$

One may exam that Greville's a_x satisfy the qualitative features. For ages 0 and 1-4 years, the Greville formula does not apply, and we use the empirical formulas of the CDW as below:

$$a_0 = \begin{cases} b_0, & m_0 \geq 0.107, \\ c_{00} + c_{01} \cdot m_0, & m_0 < 0.107, \end{cases}$$

$$a_1 = \begin{cases} b_1, & m_0 \geq 0.107, \\ c_{10} + c_{11} \cdot m_0, & m_0 < 0.107, \end{cases}$$

where the parameters are:

	b_0	b_1	c_{00}	c_{01}	c_{10}	c_{11}
Male	.33	1.352	.045	2.684	1.651	-2.816
Female	.35	1.351	.053	2.8	1.522	-1.518

(2) The oldest age group

For the oldest age group starting at age ω, traditional methods assume that the life expectancy at age ω is $1/m_{\omega+}$, where $m_{\omega+}$ is the crude death rate over age ω. This assumption requires that the population at ages over ω is stationary, which could lead to large errors. We use the m_x at ages 80-99 years to build a logistic model (Thatcher, Kannisto, and Vaupel, 1998), in which the m_x converges to 1 when x goes to infinite. This model uses he m_x at ages 80-99 years to infer that at 100 years and older, and hence is more stable than using direct data that suffer severe fluctuations caused by small population size. We stop our life tables at age 130, at which there is virtually no survivor.

(3) The variances of *LE* computed from observed and stationary deaths

When the population is stationary, Chiang's variance of estimating life expectancy at birth for an observed population, $V_O(e_0)$, reduces to variance of life expectancy of a stationary population $V_S(e_0)$, which can be indicated below using simplified formulas that neglect the details of computing the mean age of death within an age interval.

First, let the probability of surviving from birth and age i-1 to age i be l_i (with $l_0 = 1$) and p_i, life expectancy at birth can be simplified as

$$e_0 = 1 + l_1 + l_2 + ... + l_i + ..., \quad l_i = p_1 p_2 ... p_i$$

Then the (4.8, p162) of Chiang becomes

$$\frac{\partial e_0}{\partial p_i} = \frac{1}{p_i}[l_i + l_{i+1} ...] = \frac{l_i}{p_i} e_i = l_{i-1} e_i$$

Therefore, combining the (2.2, p163) and (4.10, p163) of Chiang, $V_O(e_0)$ is simplified as

$$V_O(e_0) = \sum_{i=0} [l_{i-1} e_i]^2 \frac{q_i^2 (1 - q_i)}{D_i},$$

where D_i and q_i are the number of death and the probability of death, respectively, in age interval $[i-1, i)$.

Now, let the population be stationary and the annual birth be B, there are

$$D_i = B l_{i-1} q_i,$$

and

$$V_O(e_0) = \frac{1}{B} \sum_{i=0} l_{i-1} e_i^2 q_i (1 - q_i) = \frac{1}{B} \sum_{i=0} l_i e_i^2 q_i \ .$$

We now turn to using continues version and a life-table function $T_x = \int_x^\infty l_x dx$ for concise sake. Noticing that now $q_x = -\lim_{\Delta x \to 0} \frac{l_{x+\Delta x} - l_x}{\Delta x l_x} = -\frac{1}{l_x} \frac{dl_x}{dx}$, $V_O(e_0)$ becomes

$$V_O(e_0) = -\frac{1}{B}\int_0^\infty \frac{T_x^2}{l_x^2}\frac{dl_x}{dx}dx = -\frac{1}{B}\int_0^\infty \frac{1}{l_x^2}T_x^2 dl_x = \frac{1}{B}\int_0^\infty T_x^2 d\frac{1}{l_x}$$

$$= \frac{1}{B}[-\frac{T_0^2}{l_0} - 2\int_0^\infty \frac{1}{l_x}T_x dT_x] = \frac{1}{B}[-e_0^2 + 2\int_0^\infty T_x dx]$$

$$= \frac{1}{B}[-e_0^2 - 2\int_0^\infty x dT_x] = \frac{1}{B}[-e_0^2 + 2\int_0^\infty x l_x dx] = \frac{1}{B}[-e_0^2 + \int_0^\infty l_x dx^2]$$

$$= \frac{1}{B}[-e_0^2 - \int_0^\infty x^2 dl_x] = \frac{1}{B}[-e_0^2 + \int_0^\infty x^2 \delta_x dx]$$

$$= \frac{1}{B}\int_0^\infty [x - e_0]^2 \delta_x dx = V_S(e_0).$$

$V_S(e_0)$ is the continues version of (5).

References

Agresti, A. and B. Finlay, 1997. Statistical Models for the Social Sciences. Prentice Hall, Inc., Upper Saddle River, NJ., USA, 1997.

Chiang, C. L., 1984. The Life Table and Its Applications. Robert E. Krieger Publishing Company, Malabar, FL, USA.

Coale, A.J. and P. Demeny, P., 1966. Regional Model Life Tables and Stable Populations. Princeton, N.J., Princeton University Press.

D Eayres, D. and E. S. Williams, 2004. Evaluation of methodologies for small area life expectancy estimation. Journal of Epidemiology and Community Health, (58)243–249.

Edwards, R. D. and S. Tuljapurkar, 2005. Inequality in life spans and a new perspective on mortality convergence across industrialized countries. Population and Development Review, (33) 645-674.

Greville, T. N. E., 1943. Short methods of constructing abridged life tables. The Record of the American Institute of Actuaries, (32)29-43.

Japan, Ministry of Health, Labour and Welfare, 2011. http://www.mhlw.go.jp/english/database/db-hw/lifetb10/1.html

Preston, S. H., P. Heuveline, and M. Guillot, 2001. Demography, Measuring and Modeling Population Process. Blackwell Publishers Ltd, Oxford, UK.

Silcocks, P.B.S., D. A. Jenner, and R. Reza, 2001. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. Journal of Epidemiology and Community Health, (55)38–43.

Statistics Sweden, 2011. http://www.scb.se/Pages/TableAndChart____25831.aspx.

Thatcher, A. Roger, Väinö Kannisto, and James W. Vaupel, 1998. The Force of Mortality at Ages 80 to 120. Odense, Denmark: Odense University Press.

UNDP, 2010. Human Development Report 2010. United Nations Development Programme, 1 UN Plaza, New York, NY 10017, USA.

United Nations, 2009. World Population Prospects: The 2008 Revision. United Nations, New York, 2009.

United Nations. 2011. World Population Prospects: The 2010 Revision, http://esa.un.org/unpd/wpp/index.htm.

WHO, 2005. The World Health Report 2005. World Health Organization, Geneva, Switzerland.

WHO, 2007. Civil registration: why counting births and deaths is important. http://www.who.int/mediacentre/factsheets/fs324/en/index.html.