

Using cohort change ratios to estimate life expectancy in populations with negligible migration: A new approach

David A. Swanson

Department of Sociology, University of California Riverside
david.swanson@ucr.edu

Lucky Tedrow

Department of Sociology, Western Washington University

Abstract

Census survival methods are the oldest and most widely applicable methods of estimating adult mortality, and for populations with negligible migration they can provide excellent results. The reason for this ubiquity is threefold: (1) their data requirements are minimal in that only two successive age distributions are needed; (2) the two successive age distributions are usually easily obtained from census counts; and (3) the method is straightforward in that it requires neither a great deal of judgment nor “data-fitting” techniques to implement. This ubiquity is in contrast to other methods, which require more data, as well as judgment and, often, data fitting. In this short note, the new approach we demonstrate is that life expectancy at birth can be computed by using census survival rates in combination with an identity whereby the radix of a life table is equal to 1 ($l_0 = 1.00$). We point out that our suggested method is less involved than the existing approach. We compare estimates using our approach against other estimates, and find it works reasonably well. As well as some nuances and cautions, we discuss the benefits of using this approach to estimate life expectancy, including the ability to develop estimates of average remaining life at any age. We believe that the technique is worthy of consideration for use in estimating life expectancy in populations that experience negligible migration.

Keywords: incomplete data, estimation, census data.

Résumé

Les méthodes de survie au recensement sont les méthodes les plus anciennes et les plus utilisées pour estimer la mortalité adulte et, dans le cas de populations à migration négligeables, elles peuvent fournir d'excellents résultats. Il y a trois volets à cette ubiquité : 1) les exigences de données sont minimales en ce sens que deux répartitions par âges consécutives suffisent; 2) les deux répartitions par âges successives sont habituellement faciles à obtenir à partir du recensement; 3) la méthode est claire parce qu'elle ne demande pas vraiment de jugement ni des techniques de « manipulation de données » pour être mise en œuvre. Cette ubiquité fait contraste aux autres méthodes qui ne nécessitent plus de données et de jugement ainsi que de la manipulation de données. Dans ce court article, la nouvelle approche dont nous faisons la démonstration est que l'espérance de vie à la naissance peut être calculée au moyen des taux de survie des recensements en combinaison avec une identité où la base du tableau de vie est égale à 1 ($l_0 = 1,00$). Nous soulignons que la méthode que nous proposons est moins compliquée que l'approche existante. Nous comparons les estimations en examinant notre approche par rapport aux estimations et constatons qu'elle fonctionne raisonnablement bien. Outre certaines nuances et mises en garde, nous parlons des avantages d'utiliser cette approche pour estimer l'espérance de vie, incluant la capacité de préparer des estimations de durée de vie restante à tout âge. Nous croyons que cette technique mérite considération pour l'estimation de l'espérance de vie chez les populations à migration négligeable.

Mots-clés : données incomplètes, estimation, données du recensement.

As noted in *Methods for Estimating Adult Mortality from Census Data* (United Nations 2002: 5), “Census survival methods are the oldest and most widely applicable methods of estimating adult mortality... [and can] provide excellent results [for] populations that experience negligible migration...” The reason for the ubiquity of this approach is threefold: (1) data requirements are minimal in that only two successive age distributions are needed; (2) the two successive age distributions are usually easily obtained from census counts; and (3) the method is straightforward in that it requires neither a great deal of judgment nor “data-fitting” techniques to implement. This ubiquity is in contrast to other methods, such as “Model Life Tables,” which require more data as well as judgment and, often, data fitting (United Nations 1982: 16–27). Our purpose in this paper, however, is not to debate the relative merits of these and other approaches (e.g., Swanson 1989; Swanson et al. 2009; Swanson and Palmore 1976; Swanson et al. 1977). Our purpose here is to simply demonstrate a new way of calculating life expectancy from census survival rates, one that is less involved than the existing method.¹

Census survival methods require two age distributions for a population at two points in time (generally, two successive census enumerations). Ideally, the interval between the census enumerations (e.g., 10 years) is either equal to the width of the age groups (e.g., the age groups are given in ten year increments, 0–9, 10–19...75–84, 85+) or a whole number multiple thereof (e.g., age groups given in five-year increments, 0–4, 5–9...80–84, 85+), through the final open-ended age group (e.g., 85+).

The United Nations (2002: 6) shows that using the census survival method, that the expectation of life at age x can be computed as

$$e_x = (T_x / l_{(n/2)}) / (l_x / l_{(n/2)}) = T_x / l_x \quad [1]$$

where x =age, n =the width of the age groups (up to, but not including the terminal, open-ended age group), e_x =life expectancy (average years remaining) at age x , T_x =total person years remaining to persons age x , l_x =number reaching age x , and $l_{(n/2)}$ =persons aged x to $x+n$ are assumed to be concentrated at the mid-point of the age group, and

$$l_{(x+2n/2)} / l_{(x-n/2)} = P2_{(x,n)} / P1_{(x-n,n)} \quad [2]$$

where $P2_{(x,n)}$ =number of persons counted in the second census in age group x to $x+n$ and $P1_{(x-n,n)}$ =number of persons counted in the first census in age group $x-n$ to x .

In general, then, the life-table probability of surviving from the mid-point of one age group to the next ($l_{(x+2n/2)} / l_{(x-n/2)}$) is approximated by the census survival ratio $P2_{(x,n)} / P1_{(x-n,n)}$. Continuing, the same United Nations Manual (2002: 5–6) shows that the cumulative multiplication of the probabilities shown in [2] gives the conditional survival schedule $l_x / l_{(n/2)}$. From the conditional l_x values given by [2]

1. An earlier version of this paper was presented at the 45th Actuarial Research Conference, which was held at Simon Fraser University, Burnaby, British Columbia (Canada), July 26–28, 2010. The authors are grateful for comments made by participants, including Professor Yvonne Chueh.

the conditional estimates of the number of person years lived in each age group (L_x) can be calculated as

$${}_nL_x / l_{(n/2)} = (n/2) * [(l_x / l_{(n/2)} + l_{(x+n)} / l_{(n/2)}] \quad [3]$$

where ${}_nL_x$ = number of person years lived in each age group.

Given a value of $T_x / l_{(n/2)}$ for some initial old age x , the UN shows that total remaining years expected at age x (T_x) values can be calculated as:

$$T_{(x-n)} / l_{(n/2)} = T_x / l_{(n/2)} + {}_nL_{(x-n)} / l_{(n/2)} \quad [4]$$

This leads us back to equation [1], so that the expectation of life at age x using the United Nations (2002) approach is:

$$e_x = (T_x / l_{(n/2)}) / (l_x / l_{(n/2)}) = T_x / l_x$$

In our approach, we start by noting that when the radix of a life table is equal to 1 ($l_0 = 1.00$) life expectancy at birth can be computed directly from the expression

$$e_0 = S_0 + (S_0 * S_1) + (S_0 * S_1 * S_2) + \dots + (S_0 * S_1 * S_2 \dots * S_x) \quad [5]$$

where e_0 = life expectancy at birth, S_0 = survivorship from $t=0$ (e.g., birth) to $t=1$ (e.g., age 1), S_1 = survivorship from $t=1$ (e.g., age 1) to $t=2$ (e.g., age 2), and so on through S_x , and $S_x = {}_1L_x / {}_1L_{(x-n)}$.

Equation [5] is set up for single year age groups. However, we can generalize it to other age groups: $S_x = {}_nL_x / {}_nL_{(x-n)}$, so that

$$e_0 = {}_nS_0 + ({}_nS_0 * {}_nS_1) + ({}_nS_0 * {}_nS_1 * {}_nS_2) + \dots + ({}_nS_0 * {}_nS_1 * {}_nS_2 \dots * {}_nS_x) \quad [5a]$$

As equation [5] and equation [5a] both imply, the fundamental life table function is inherent in our method. That is, via the ${}_nS_x$ values, we have ${}_nq_x$ values. The Appendix shows a derivation of the relationship between survivorship rates and life expectancy, as shown in equation [5] and generalized to equation [5a].

We also apply census survival rates, although we prefer to use the more general term “cohort change ratios” (CCRs). CCRs enter into a range of measures and applications, such as the “Hamilton-Perry Method,” which is often used for doing short-term population forecasts of small areas such census tracts (Smith et al. 2001: 153–8). Following Smith, Tayman, and Swanson (2001: 155) and using notation from equation [2], a CCR can be generally defined as

$${}_nCCR_x = P2_{(x,n)} / P1_{(x-n,n)} \quad [6]$$

Our new approach is the result of combining, on the one hand, either equation [5] or [5a] for computing life expectancy with, on the other hand, equation [6] in order to estimate e_x . Starting with

$${}_nS_x = {}_nL_x / {}_nL_{(x-n)} \approx P2_{(x,n)} / P1_{(x-n,n)} \quad [7]$$

we have, as shown in equation [5a],

$$e_0 = {}_nS_0 + ({}_nS_0 * {}_nS_1) + ({}_nS_0 * {}_nS_1 * {}_nS_2) + \dots + ({}_nS_0 * {}_nS_1 * {}_nS_2, \dots, * {}_nS_x)$$

As is the case with the more involved United Nations (2002) approach, this approach will only work for populations for which migration is negligible, but there are many areas of interest around the world where this is the case, or approximately so (United Nations 2002). The world as a whole meets this requirement. Countries with negligible migration include North Korea and Burma, among others. Other such populations are found in the historical record—the former Soviet Union, Albania from 1950 to 1980, and the Peoples Republic of China from 1950 through 1970, for example. Still others may be defined by race and ethnicity or other ‘rules’ of membership (e.g., Indigenous Populations in Australia and Canada, Native Hawaiians; native-born populations).

Broadly speaking, the method can be applied to any population subject to renewal through a single increment (birth) and extinction through a single decrement (death), where there are at least two successive census counts that provide the population by age. We also note that unlike the UN method, the approach we take can be used to yield estimates of life expectancy at birth. Moreover, like the UN approach, this one is not subject to the limitations imposed by stationary or even stable population requirements.

We have developed life expectancy estimates directly from cohort change ratios constructed for the world as a whole and Burma. We compare our life expectancy estimates to estimates for the world as a whole and Burma for the period 1950–55 to 2045–50 (Census Bureau 2010). The comparison of our estimates of e_0 for the world as a whole and those of the United Nations are found in Table 1. The Mean Absolute Percent Difference between our estimates and those made by the United Nations over the entire period is 3.28 per cent, while the Mean Algebraic Percent Difference is 0.68 per cent. Both of the summary measures indicate rather close agreement between the two sets of estimates, and the latter suggests that our estimates are only slightly upwardly biased relative to those of the United Nations.

Table 2 shows e_0 estimates for Burma for the period 1975–80 to 2005–10. As the table shows, all of our e_0 estimates are less than those done by the U.S. Census Bureau, which means the mean Algebraic Percent Difference of –4.50 per cent is the same as the Mean Absolute Percent Difference. Our estimates of life expectancy at birth remain almost constant at age 60 from 1995–2000 to 2005–10, while those available from the Census Bureau increase from 61 to 65 years. However, we again find that the summary measures suggest reasonably close agreement.

As well as some nuances (e.g., converting CCRs into survival ratios may require additional refinements) and cautions (e.g., the population data by age may be faulty), we find benefits in using this approach to estimate life expectancy, including the ability to develop estimates of average remaining life at any age (not shown here). We suggest that the technique is worthy of consideration for use in estimating life expectancy in populations experiencing negligible migration, given the cautions we discuss. As such, we believe that this new approach adds another dimension to census survival methods—which, as we noted at the outset, are “... the oldest and most widely applicable methods of estimating adult mortality...

Table 1a. Comparison of life expectancy estimates for the world as whole calculated from cohort change ratios (during each period), 1950–55 to 2045–50 shown in equation [7] and equation [5a] with estimated values available from the US Census Bureau (2010).

YEARS	e_0 calculated from UN Census S_x values using [6] and [4]*	e_0 calculated from UN Census S_x values using [6] and [4]*	e_0 calculated from UN Census S_x values using [6] and [4]*	UN Life Expectancy(e_0)		
	total e_0	male e_0	female e_0	total e_0	male e_0	female e_0
1950–55	51.80	50.28	53.38	46.6	45.2	48
1955–60	53.95	52.01	56.01	49.5	48.1	50.9
1960–65	56.00	54.61	57.35	52.4	51	53.7
1965–70	58.43	56.69	60.15	56.1	54.6	57.6
1970–75	60.40	58.89	61.85	58.2	56.6	59.8
1975–80	61.77	59.73	63.81	60.2	58.3	62
1980–85	62.61	60.48	64.77	61.7	59.7	63.7
1985–90	63.49	61.27	65.76	63.2	61.2	65.2
1990–95	64.05	61.55	66.65	64	61.9	66.2
1995–2000	64.90	62.39	67.54	65.2	63	67.4
2000–05	65.80	63.35	68.34	66.4	64.2	68.6
2005–10	66.65	64.30	69.07	67.6	65.4	69.8
2010–15	67.65	65.42	69.93	68.9	66.7	71.1
2015–20	68.57	66.36	70.83	70.1	67.9	72.3
2020–25	69.38	67.17	71.65	71.1	68.9	73.4
2025–30	70.14	67.91	72.42	72.1	69.9	74.4
2030–35	70.81	68.58	73.09	73.1	70.8	75.4
2035–40	71.43	69.19	73.72	73.9	71.7	76.3
2040–45	72.02	69.79	74.31	74.8	72.5	77.1
2045–50	72.59	70.37	74.88	75.5	73.3	77.9

The calculations are found in a worksheet available from the authors.

* e_0 is estimated by using equation [5a].

Table 1b. Differences between life expectancy estimates of the total population calculated using cohort survival ratios and those produced by the U.S. Census Bureau, 1950–55 to 2045–50.

Years	Total e_0 estimated via CCR method	Total e_0 estimated by U.S. Census Bureau	Difference	% difference	Absolute % difference
1950–55	51.80	46.60	5.20	11.17%	11.17%
1955–60	53.95	49.50	4.45	8.99%	8.99%
1960–65	56.00	52.40	3.60	6.87%	6.87%
1965–70	58.43	56.10	2.33	4.15%	4.15%
1970–75	60.40	58.20	2.20	3.77%	3.77%
1975–80	61.77	60.20	1.57	2.60%	2.60%
1980–85	62.61	61.70	0.91	1.47%	1.47%
1985–90	63.49	63.20	0.29	0.46%	0.46%
1990–95	64.05	64.00	0.05	0.08%	0.08%
1995–2000	64.90	65.20	–0.30	–0.45%	0.45%
2000–05	65.80	66.40	–0.60	–0.91%	0.91%
2005–10	66.65	67.60	–0.95	–1.41%	1.41%
2010–15	67.65	68.90	–1.25	–1.82%	1.82%
2015–20	68.57	70.10	–1.53	–2.19%	2.19%
2020–25	69.38	71.10	–1.72	–2.41%	2.41%
2025–30	70.14	72.10	–1.96	–2.72%	2.72%
2030–35	70.81	73.10	–2.29	–3.14%	3.14%
2035–40	71.43	73.90	–2.47	–3.34%	3.34%
2040–45	72.02	74.80	–2.78	–3.72%	3.72%
2045–50	72.59	75.50	–2.91	–3.85%	3.85%
			MALPE	0.68%	
			MAPE	3.28%	

Table 2. Comparison of life expectancy estimates for Burma calculated from cohort change ratios (during each period), 1975–80 to 2005–10 as shown in equation [7] and equation[5a] with estimated values available from the US Census Bureau (2010).

Source of estimate/ year	Estimated life expectancy from CCRs	Life expectancy at birth (US Census)	Difference	% Difference	Absolute % Difference
1975–80	49.99	54	–4.01	–7.43%	7.43%
1980–85	52.47	56	–3.53	–6.31%	6.31%
1985–90	55.96	56	–0.04	–0.08%	0.08%
1990–95	56.97	59	–2.03	–3.43%	3.43%
1995–2000	60.09	61	–0.91	–1.50%	1.50%
2000–05	60.82	63	–2.18	–3.46%	3.46%
2005–10	60.99	65	–4.01	–6.17%	6.17%
			MALPE	–4.05%	
			MAPE	4.05%	

(and can) provide excellent results (for) populations that experience negligible migration..." (United Nations 2002: 5).

References

- Smith, S.K., J. Tayman, and D.A. Swanson. 2001. *State and Local Population Projections: Methodology and Analysis*. New York, NY: Kluwer Academic/Plenum Publishers.
- Swanson, D.A. 1989. A state-based regression model for estimating substate life expectancy. *Demography* 26(February):161–70.
- Swanson, D.A., and J.A. Palmore. 1976. Two parameter regression estimates of current life expectancy at birth: Part I. *Asian & Pacific Census Forum* 3(November):5–10.
- Swanson, D.A., J.A. Plamore, and C. Sundaram. 1977. Two parameter regression estimates of current life expectancy at birth: Part II. *Asian & Pacific Census Forum* 5(May):5–10.
- Swanson, D.A., M. McGehee, and N. Hoque. 2009. Socio-economic status and life expectancy in the United States, 1970–1990. *Population Review* 48(1):39–63.
- United Nations. 1982. *Model Life Tables for Developing Countries*. New York, NY: Population Division, United Nations.
- . 2002. *Methods for Estimating Adult Mortality*. New York, NY: Population Division, United Nations.
- United States Census Bureau. 2010. World Population Statistics, available at <http://www.census.gov/ipc/www/idb/informationGateway.php>.

Appendix. Relation between survival rates and life expectancy

Any particular set of age-specific survival rates implies a specific life expectancy. As an example, using a complete single year life table (where x =a given age), the relationship between a set of survival rates S_x and L_x , the corresponding entries in the “years lived” column of the life table, is $S_x = L_x / L_{x-1}$ for ages 1 and over, while $S_0 = L_0$ for survivors from birth to the age zero age group. Thus, in a life table with a radix=1.0, life expectancy at birth can be expressed as $e_0 = L_1 + L_2 + L_3 + \dots + L_x$. That is, life expectancy can be expressed as the sum of the L_x values. It is readily seen that

$$L_0 = S_0, L_1 = S_0 * S_1, L_2 = S_0 * S_1 * S_2 \dots L_x = S_0 * S_1 * S_2 \dots S_{x-1} * S_x$$

Substituting S_x for L_x in the preceding yields Equation [5]:

$$e_0 = S_0 + (S_0 * S_1) + (S_0 * S_1 * S_2) + \dots + (S_0 * S_1 * S_2 * \dots * S_x)$$

and Equation [5] can be generalized to apply to an abridged life table and expressed as Equation[5a]

$$e_0 = {}_nS_0 + ({}_nS_0 * {}_nS_1) + ({}_nS_0 * {}_nS_1 * {}_nS_2) + \dots + ({}_nS_0 * {}_nS_1 * {}_nS_2 * \dots * {}_nS_x)$$