The Stata Journal (2023)
23, Number 3, pp. 813–834 DOI: 10.1177/1536867X231196441

Iterative intercensal single-decrement life tables using Stata

Jerônimo Oliveira Muniz
Universidade Federal de Minas Gerais
Belo Horizonte, Brazil
jeronimo@ufmg.br

Abstract. One way to estimate mortality in countries with incomplete data is to utilize intercensal methods, which do not require model life tables and provide accurate results even in the presence of age distortions and death underregistration. In this article, I revisit three of these techniques (census based, death distribution, and an iterative procedure) and introduce `ilt`, a command to calculate single-decrement life tables and the net flow of migrants by age. The required inputs are two age-specific population distributions and the average number of deaths between them. The empirical example draws on data from Vietnam, but the methods are extendable to any context and period.

Keywords: st0729, ilt, age, demography, life expectancy, life table, iterative, intercensal, census

1 Introduction

Period life tables describe the mortality experience of individuals over their life cycle. They provide death rates, survival probabilities, and life expectancies by age. Actuaries and demographers often use them to project populations. Such tables may have single or multiple sources of decrement. "A single decrement process is one in which individuals have only one recognized mode of exit from a defined state. Mortality is one such process" (Preston, Heuveline, and Guillot 2001). It differs from a multiple-decrement process, in which individuals have more than one mode of exit, and from multistate life tables, in which "increments and decrements from more than one state are allowed" (Muniz 2020, 721). In this article, I implement and compare direct methods for constructing single-decrement life tables. Therefore, I assume that mortality is the only source of attrition that affects the life experience of a cohort.

In the early 1980s, four articles demonstrated how to estimate mortality using successive age distributions and intercensal deaths (Bennett and Horiuchi 1981, 1984; Preston and Bennett 1983; Preston and Coale 1982). In the 1990s, this procedure was improved for accuracy and to facilitate its applicability (Merli 1998; Preston et al. 1996). Intercensal methods do not depend on model life tables (as indirect mortality methods do), bypass the stability assumption, and are sufficiently accurate even when death registration is incomplete. Developing countries with deficient death registration benefit from intercensal techniques because they allow the estimation of life tables directly from census age distributions and age-specific growth rates (Preston and Bennett 1983;

Spoorenberg 2008). In addition, these rates can be used to adjust the reported number of deaths and estimate life tables even in populations with fairly accurate age reporting (Bennett and Horiuchi 1984; Merli 1998; Preston et al. 1996).

In this article, I demonstrate the use of `ilt`, a command for estimating single-decrement life tables using three intercensal mortality estimation methods: census based, death distribution, and an iterative procedure that reconciles the two previous techniques. The program also estimates net migration flows by comparing observed and projected populations. The following section reviews these methods and illustrates their application using data from Vietnam.

2 Required data for intercensal methods

Preston and Bennett (1983) introduced the census-based method of estimating single-decrement life tables using two successive age distributions. A year later, Bennett and Horiuchi (1984) developed the death-distribution procedure, which was improved by Preston et al. (1996) to estimate mortality from a set of recorded deaths by age and age-specific growth rates. And more recently, Merli (1998) proposed an iterative mechanism between these two methods to overcome some of the limitations in both procedures and to reconcile their results. To ensure analytical comparability, I replicate the results of Merli (1998) using male data from the 1979 and 1989 Vietnamese censuses.[1]

These direct methods for estimating single-decrement life tables rely on several assumptions. Therefore, before applying them, the user must ensure data consistency, namely, that 1) populations are closed to migration or have few migration flows, 2) census coverage and completeness of death reporting are the same for all ages, and 3) there is no substantial misreporting of age and sex (Queiroz et al. 2020, 3). In particular, residual migration is a source of error that potentially biases the estimates. For this reason, as will become clear in subsection 2.3, the iterative method attempts to remove the influence of residual migration by forward projecting the baseline population and using this second projected or corrected distribution to compute the life table describing the life experience of a cohort.

1. Population counts were "adjusted for the effects of intercensal emigration on an age- and sex-specific basis" (Merli 1998, 348–349, app. 1). The results shown are slightly different from those in Merli (1998, 351) because I have restricted the final age group to 80 and over instead of 85.

The appropriate data structure to use before applying these methods is[2]

```
. use pop_vietnam
. list, noobs separator(17)
```

age_x	N_t1	N_t2	D
0	3946224	4710423	48508
5	3928795	4430179	8029
10	3632555	3898298	3928
15	2954333	3427357	3783
20	2281171	2974282	3856
25	1742277	2832160	3469
30	1177320	2361692	3053
35	966580	1604918	3093
40	919291	1048246	3345
45	994602	877589	4836
50	825356	866821	6215
55	680996	918363	9138
60	540920	725079	12070
65	419164	533445	13645
70	284003	329167	14310
75	183222	215510	14357
80	103773	145637	14787

The first column must always define the lower bound of the age groups. The remaining columns are the number of individuals (N_x) and the average annual number of deaths (D_x) between time $t1$ and $t2$ at age x. After entering a dataset with this structure and these variables into Stata, we are ready to explore the full potential of `ilt`.

2.1 Census-based method

This method, also known as the r-method, requires two successive age distributions and a set of derived age-specific growth rates to measure mortality conditions. The method does not use model life tables and does not assume population stability. It does not depend on the distribution of deaths (D_x) but is sensitive to age distortions and intercensal net migration. In this method, the amount of population exposed to the risk of the event, or the number of person-years L_x, is[3]

$$L_x = N_x^* \exp(S_x)$$

where

$$N_x^* = \sqrt{N_x(t1)N_x(t2)}$$

2. I used `texdoc` to import Stata outputs into LaTeX (Jann 2016).

3. I present demographic equations in their discrete forms to reflect their common use for five-year age intervals.

is the geometric mean of the population at age x between times $t1$ and $t2$, and

$$S_x = \frac{n}{2} r_x + n \sum_{a=0}^{x-5} r_a$$

is the cumulation of age-specific growth rates to the midpoint of age interval of length n, where

$$r_x(t1, t2) = \frac{\ln \left\{ \frac{N_x(t2)}{N_x(t1)} \right\}}{t2 - t1} \tag{1}$$

refers to age-specific growth rates between times $t1$ and $t2$. The remaining functions of the life table are the number of individuals in the cohort surviving to age x (l_x), the number of person-years lived above age x (T_x), and the life expectancy of individuals at age x (e_x).[4] These functions are

$$l_x = \frac{L_x + L_{x-5}}{2n} \; \forall x > 0$$

$$T_x = \sum_{a \geq x}^{\infty} L_a \tag{2}$$

$$e_x = \frac{T_x}{l_x} \; \forall x > 0 \tag{3}$$

In Stata, the census-based life table for Vietnamese males is obtained by entering the following command line:

```
. ilt, interval(10) census
Table 1. Census-based method
```

age x	r	S	L	T	l	e
0	.0177019	.0442546	4506510	5.41e+07	.	.
5	.0120107	.1185361	4696993	4.96e+07	920350.3	53.89616
10	.0070604	.1662138	4443543	4.49e+07	914053.5	49.1288
15	.0148517	.2209939	3969043	4.05e+07	841258.6	48.09795
20	.0265314	.3244515	3603110	3.65e+07	757215.3	48.19471
25	.0485847	.5122416	3707500	3.29e+07	731061	44.99031
30	.0696138	.8077377	3739856	2.92e+07	744735.7	39.18593
35	.0507064	1.108538	3773786	2.54e+07	751364.3	33.86281
40	.0131271	1.268122	3488963	2.17e+07	726275	29.83652
45	-.0125164	1.269648	3325612	1.82e+07	681457.5	26.67893
50	.0049018	1.250612	2954059	1.49e+07	627967.1	23.65561
55	.0299036	1.337625	3013025	1.19e+07	596708.4	19.94422
60	.0293009	1.485637	2766706	8887860	577973.1	15.37764
65	.0241094	1.619162	2387427	6121154	515413.3	11.87621
70	.014758	1.716331	1701229	3733727	408865.6	9.13192
75	.0162309	1.793803	1194707	2032499	289593.5	7.018454
80	.0338911	1.919108	837791.9	837791.9	203249.9	4.12198

Because the input populations (`N_t1` and `N_t2`) are 10 years apart, I must specify `interval(10)`. I omitted the age subscript x from the column names to avoid clutter and stored the columns as Mata vectors of the same name (`r`, `S`, `L`, `T`, `l`, `e`). Life

4. `ilt` calculates T_x using `flipud()`, a function for flipping matrices (Baum 2006).

expectancies at ages 5, 10, and 15 are 53.90, 49.13, and 48.10 years, respectively. According to Merli (1998, 350), these values are underestimated because of undercounting in the young and mid-adult ages or distortions in the 1989 age structure due to residual emigration. The following method incorporates the age distribution of intercensal deaths to improve these estimates.

2.2 Death-distribution method

Bennett and Horiuchi (1984) developed a technique for estimating mortality when there is undercounting in the number of deaths and inaccuracy in age reporting. The method calculates the number of deaths in the life table (d_x) from the average number of annual deaths recorded in the population (D_x) between $t1$ and $t2$. The method converts the distribution of reported deaths in the population to the corresponding distribution of deaths in the life table using intercensal age-specific growth rates. It assumes that underregistration is constant by age.

Preston et al. (1996) used the death-distribution method to estimate mortality rates for African Americans over the age of 64. Merli (1998) used it to determine mortality conditions in Vietnam, and Hill (2001) and Castanheira and Monteiro da Silva (2022) used it to assess the relative completeness of death counts relative to censuses. According to Preston et al. (1996), in the discrete case, the number of deaths between ages x and $x+n$ in the life table for five-year age groups is a function of the number of recorded deaths and the cumulation of age-specific growth rates:

$$d_x = D_x \exp\left\{\frac{n}{2}\left(\sum_{a=0}^{x} r_a + \sum_{a=5}^{x-5} r_a\right)\right\} \tag{4}$$

Equation (4) assumes that $d_0 = D_0$ and $d_5 = D_5 \exp\{(n/2)(r_0 + r_5)\}$. This gives the same results as the equation described in Preston, Heuveline, and Guillot (2001, 188):

$$d_x = d_{x-5}\frac{D_x}{D_{x-5}} \exp\left\{\frac{n}{2}\left(r_{x-5} + r_x\right)\right\} \tag{5}$$

However, (5) highlights that the death-distribution method assumes that all recorded deaths are deficient in the same proportion, leaving the ratio $D_x \backslash D_{x-5}$ undisturbed (Preston, Heuveline, and Guillot 2001, 179). Violating this assumption will bias life expectancy estimates: “Completeness of death registration declining with age distorts the $D_x \backslash D_{x-5}$ ratio [...] and introduces a downward bias in the estimates of life expectancy at all ages below the ages at which deaths are omitted” (Merli 1998, 352). Moreover, the method is not immune to differential completeness of census coverage and intercensal migration, because it uses age-specific growth rates between two population distributions. However, the death-distribution method is less sensitive to these problems than the census-based method and has the advantage of providing an estimate of the life expectancy at birth.[5]

5. Bennett and Horiuchi (1984, 222) show that life expectancy at age 5 is 10 times more sensitive to errors in the growth rate in the census-based method than in the death-distribution method.

To continue the life table, I estimate the number of person-years lived by the cohort between ages x and $x+n$ (L_x) as[6]

$$L_x = nl_{x+n} + \frac{n}{2} d_x \; \forall x > 0 \quad \text{and} \quad x \leq 75$$

The probability of dying between age x and $x+n$ (q_x) is equal to the ratio of deaths to survivors:

$$q_x = \frac{d_x}{l_x} = \frac{l_x - l_{x+n}}{l_x}$$

I calculate T_x and e_x as in (2) and (3). The mortality estimates, based on the death-distribution method, are accessed in Stata with the following command:

```
. ilt, interval(10) death
```

Table 2. Death-distribution method

age x	r	d	Ld	Td	ld	qd
0	.0177019	48508	2526667	3.31e+07	546274.2	.0887979
5	.0120107	8648.115	2467211	3.06e+07	497766.2	.0173739
10	.0070604	4437.493	2434497	2.81e+07	489118.1	.0090724
15	.0148517	4514.329	2412117	2.56e+07	484680.6	.009314
20	.0265314	5102.993	2388074	2.32e+07	480166.2	.0106276
25	.0485847	5539.218	2361468	2.08e+07	475063.2	.01166
30	.0696138	6550.935	2331243	1.85e+07	469524	.0139523
35	.0507064	8965.868	2292451	1.62e+07	462973.1	.0193659
40	.0131271	11374.04	2241601	1.39e+07	454007.2	.0250526
45	-.0125164	16469.03	2171993	1.16e+07	442633.2	.037207
50	.0049018	20766.12	2078905	9450546	426164.1	.048728
55	.0299036	33308.49	1943719	7371641	405398	.0821624
60	.0293009	51014.25	1732912	5427922	372089.5	.1371021
65	.0241094	65909.39	1440603	3695010	321075.3	.2052771
70	.014758	76175.12	1085392	2254407	255165.9	.2985317
75	.0162309	82581.54	688500.1	1169015	178990.8	.4613732
80	.0338911	96409.24	480515.1	480515.1	96409.24	1

6. For the first age group, $L_0 = nl_5 + 0.78 d_0$ (see Preston, Heuveline, and Guillot [2001, 188]). However, life expectancies are not sensitive to the proportion of deaths at age x (the separation factor at age x). In the last age group, the number of person-years is ${}_\infty L_{80} = l_{80} \log_{10}(l_{80})$ (see Ortega [1982, 25] and United Nations [1956, 23]). Using these equations instead of those based on the West model life table (Merli 1998, 352) is inconsequential for the Vietnamese mortality standards and has the added virtue of generalizing the program to a broader class of countries and contexts. As usual, there may be differences between estimates because of differences in data sources, methods, and assumptions.

age x	ed
0	60.55177
5	61.37661
10	57.41761
15	52.92041
20	48.39444
25	43.88742
30	39.37569
35	34.89747
40	30.53727
45	26.25772
50	22.17584
55	18.18371
60	14.58768
65	11.50823
70	8.835063
75	6.531147
80	4.984119

The life expectancies shown in column `ed` are less sensitive to the age-specific growth rates derived from the Vietnamese age structure. Life expectancies at ages 5, 10, and 15 are now higher than in the census-based method and are 61.38, 57.42, and 52.92 years, respectively. In addition to the number of deaths in the cohort (d_x), `ilt` calculates the probability of dying (q_x). For example, the probability of dying (column `qd`) between the ages of 40 and 45 is 0.025 or 2.5%. Except for the first two age groups, the values in the `qd` column increase with age, as they should.

2.3 The iterative procedure

To avoid the bias implicit in the growth rates associated with age distortions, differential coverage, and residual intercensal migration, Giovanna Merli (1998) developed a two-step iterative procedure to reconcile the census-based and death-distribution methods.

In the first step, I project the population forward using life table survival probabilities computed using the Bennett and Horiuchi (1984) death-distribution method. Survivorship ratios (p_x) between ages x and $x+n$ represent the proportion surviving the intercensal period. In the case of 5 years, they are

$$p_x = \frac{L_{x+5}}{L_x} \; \forall x < 75 \tag{6}$$

And for the next-to-last open age group, at age 75, the survivorship ratio is

$$p_{75} = \frac{L_{80}}{L_{75} + L_{80}} \tag{7}$$

Multiplying these ratios by the 1979 base population projects the Vietnamese male population forward to 1984. Multiplying the 1984 projected population again by its age-specific survivorship ratio gives the projected population for 1989, in $t2$. Thus,

$$N_{x+10}(t2)^{\text{proj}} = N_x(t1)p_x p_{x+5} \qquad \text{for} \quad x \geq 10 \quad \text{and} \quad x \leq 75 \tag{8}$$

Equation (8) results on the number of people at age $x+10$ years later. For the first age group, the projected population in 1984 uses the third equation in Merli (1998, 356, footnote 6). For example,

$$N_0(1984)^{\text{proj}} = N_0(1979)p_0 \exp\left\{-\frac{n}{2}\left(r_0 + r_5\right)\frac{L_0}{L_5}\right\}$$

And for the last age group,

$$N_{80}(1984)^{\text{proj}} = N_{75}(1979)p_{75} + N_{80}(1979)p_{75}$$

The resulting survivorship ratios (p_x), described in (6) and (7), and the population projected to time $t2$—using p_x derived from the death-distribution method—are stored in Mata and can be retrieved by typing

```
. mata p, N_proj
                  1             2
     ┌─────────────────────────────┐
   1 │  .9764683966   3231641.888  │
   2 │  .9867404828    3487070.06  │
   3 │  .9908073168   3802269.289  │
   4 │  .9900322807    3841063.79  │
   5 │  .9888590008   3563286.635  │
   6 │  .9872005971   2892298.896  │
   7 │   .983359945    2226884.14  │
   8 │  .9778185954   1691356.385  │
   9 │  .9689473349   1132049.268  │
  10 │  .9571417282   915790.7852  │
  11 │  .9349722635   852568.7918  │
  12 │  .8915445847   890070.2926  │
  13 │  .8313191387   687991.5538  │
  14 │  .7534287568   504725.6853  │
  15 │   .634333255    338799.695  │
  16 │  .4110426554   200328.9194  │
  17 │            0   122539.9272  │
     └─────────────────────────────┘
```

Column 1 shows the proportion of people surviving the intercensal period (p_x) between ages x and $x+n$. For example, 97.65% of those persons at age 0 (in row 1) could expect to be alive five years later in a stationary population subject to the mortality conditions of the above-stated life table. Multiplying these survivorship ratios by the male Vietnamese population in 1979 gives the vector shown in column 2, which is the projected population in 1989, $N_{x+10}(t2)^{\text{proj}}$. In 1979, there were 3,946,224 individuals alive at age 0. Ten years later, in 1989, we should expect to have 3,802,269 individuals alive at age 10 (in row 3), assuming the mortality conditions of the life table based on the death-distribution method.

As a result of this exercise, we obtain two age distributions. The first represents the 1979 population of Vietnam, and the second represents the population projected 10 years later, which is less affected by coverage errors and intercensal migration than the original 1989 census population. From these two distributions (observed population in $t1$ and projected population in $t2$), we compute a new set of age-specific growth rates and a new life table using the Preston and Bennett (1983) census-based model. The difference is that while we use the projected population to estimate age-specific growth rates, we simultaneously correct for the sensitivity of the census-based method to differential completeness of census enumeration and residual migration.[7] Moreover, estimates obtained by applying forward (or backward) projections are proven more reliable than other intercensal techniques based on growth rates and age distributions of deaths (Palloni and Kominski 1984, 492).

In the second stage of the iterative process, to correct the initial growth rates estimated in the death-distribution procedure, I conduct a new round of forward projections but this time using the new intercensal growth rates estimated in the first stage of the process using the census-based method. Figure 1 diagrams the iterative cycle in four steps.

7. Racial reclassification could also account for migration in populations disaggregated by race, or it could account for social mobility in a population segmented by income, wealth, education, or occupation.

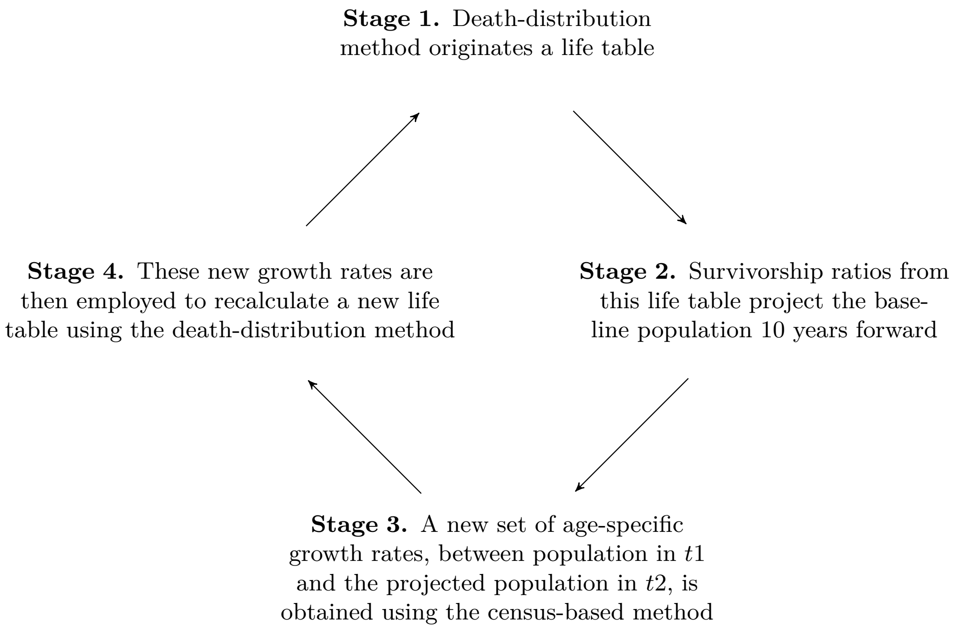

Figure 1. Iterative procedure between census-based and death-distribution methods

The iteration process between the census-based and the death-distribution methods continues as long as the difference—see [M-5] **reldif( )**—between the growth rates in the previous and the next iteration cycle is greater than 1e−10. In the case of Vietnamese males, the program reaches convergence after 68 iterations. To obtain life tables based on the iterative procedure, one must type

```
. ilt, interval(10) iterative
```

Table 3. Death-distribution method with corrected growth rates (iterative method)

age x	ri	di	Li	Ti	li	qi
0	-.0021469	48508	2432033	3.15e+07	527347.4	.0919849
5	-.0030298	7925.761	2374383	2.91e+07	478839.4	.016552
10	.0046311	3893.046	2344836	2.67e+07	470913.7	.008267
15	.0264031	4051.812	2324974	2.44e+07	467020.6	.0086759
20	.0447032	4933.472	2302510	2.20e+07	462968.8	.0106562
25	.0506832	5633.579	2276093	1.97e+07	458035.3	.0122994
30	.0636382	6598.265	2245513	1.75e+07	452401.8	.014585
35	.0558127	9011.042	2206490	1.52e+07	445803.5	.020213
40	.0206004	11796.57	2154471	1.30e+07	436792.4	.0270073
45	-.0087106	17569.33	2081056	1.08e+07	424995.9	.04134
50	.002413	22226.57	1981566	8763920	407426.6	.0545536
55	.0254806	35040.28	1838399	6782353	385200	.0909665
60	.0221368	52134.31	1620463	4943954	350159.7	.1488872
65	.0160103	64834.7	1328040	3323491	298025.4	.2175476
70	.0143962	73364.7	982541.7	1995451	233190.7	.3146125
75	.004346	77136.57	606288.5	1012910	159826	.4826285
80	.0116553	82689.42	406621.1	406621.1	82689.42	1

age x	ei
0	59.74482
5	60.71815
10	56.698
15	52.14979
20	47.58431
25	43.06991
30	38.57511
35	34.10906
40	29.76115
45	25.51784
50	21.51043
55	17.60736
60	14.11914
65	11.15171
70	8.557165
75	6.337577
80	4.91745

Table 4. Census-based method with corrected growth rates (iterative method)

age x	ri	Sc	Lc	Tc	lc	ec
0	-.0021469	-.0053671	3883193	5.02e+07	.	.
5	-.0030298	-.0183088	3799521	4.63e+07	768271.4	60.3191
10	.0046311	-.0143057	3664844	4.25e+07	746436.5	56.99335
15	.0264031	.0632797	3591493	3.89e+07	725633.8	53.57671
20	.0447032	.2410455	3630064	3.53e+07	722155.8	48.86145
25	.0506832	.4795114	3625964	3.17e+07	725602.8	43.6265
30	.0636382	.7653149	3479006	2.80e+07	710497	39.45062
35	.0558127	1.063942	3702522	2.46e+07	718152.9	34.18568
40	.0206004	1.254975	3574497	2.08e+07	727701.9	28.64912
45	-.0087106	1.284699	3440903	1.73e+07	701540	24.62229
50	.002413	1.268956	2971540	1.38e+07	641244.2	21.57153
55	.0254806	1.33869	2950261	1.09e+07	592180.1	18.34084
60	.0221368	1.457733	2595899	7910819	554616	14.2636
65	.0160103	1.553101	2146115	5314920	474201.4	11.20815
70	.0143962	1.629117	1556326	3168805	370244.1	8.55869
75	.004346	1.675973	1000646	1612479	255697.2	6.306204
80	.0116553	1.715976	611832.5	611832.5	161247.9	3.79436

Note: Convergence achieved after 68 iterations.

The iterative algorithm approximates the life expectancies provided by the census-based and death-distribution methods. At the end of the iterative process, we have two new life tables. The first matrix (`Table_3`) uses the death-distribution method, but this time with corrected age-specific growth rates from the iterative process. The second matrix (`Table_4`) shows improved results for the census-based procedure, accounting for differential completeness of census enumeration and residual emigration.

Last, the difference between observed and projected populations at $t2$ (1989)—using survivorship ratios derived from the death-distribution method with corrected growth rates—accounts for residual migration (United Nations 1970). Including the `residual` option on the `ilt` command line shows the observed and projected populations at $t2$, the difference, and the ratio between them.

```
. ilt, interval(10) residual
```

Table 5. Difference and ratio btw. observed and projected populations

age x	N(t2)	N(t2)proj	N(t2)-N(t2)proj	N(t2)/N(t2)proj
0	4710423.000	3862407.078	848015.922	1.220
5	4430179.000	3811545.282	618633.718	1.162
10	3898298.000	3804736.839	93561.161	1.025
15	3427357.000	3847039.659	-419682.659	0.891
20	2974282.000	3566985.724	-592703.724	0.834
25	2832160.000	2892220.350	-60060.350	0.979
30	2361692.000	2224702.007	136989.993	1.062
35	1604918.000	1688998.198	-84080.198	0.950
40	1048246.000	1129586.626	-81340.626	0.928
45	877589.000	911632.196	-34043.196	0.963
50	866821.000	845514.400	21306.600	1.025
55	918363.000	878628.667	39734.333	1.045
60	725079.000	674950.224	50128.776	1.074
65	533445.000	491944.333	41500.667	1.084
70	329167.000	327978.204	1188.796	1.004
75	215510.000	191360.413	24149.587	1.126
80	145637.000	116601.138	29035.862	1.249

Note: N(t2)proj uses survivorship ratios derived from Table 3.

Because of emigration, the observed population between the ages of 15 and 25 at time $t2$ is about one million people less than the projected population. Matrix `Table_5` also shows that the youngest and oldest age groups had the largest inflows of migrants, because the ratios between observed and projected populations $\{N_x(t2)/N_x(t2)^{\text{proj}}\}$ are the largest in these age groups. To generate comparative graphs of life expectancies (figure 2), projected populations (figure 3), and ratios between observed and estimated projected populations by age (figure 4), one must enter options `life`, `proj`, and `ratio`, respectively.

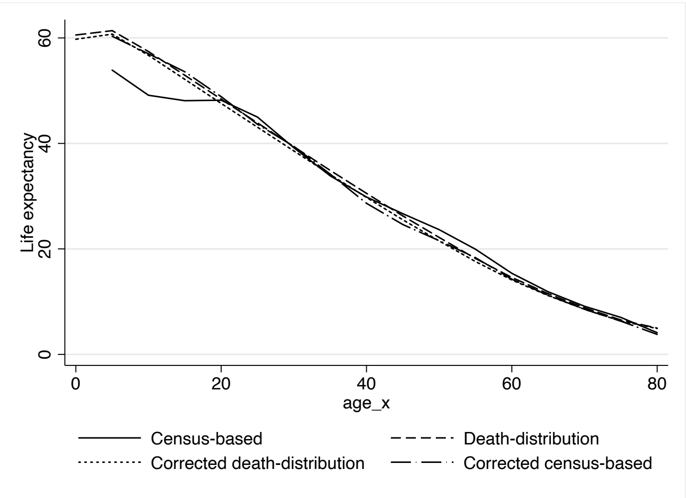

Figure 2. Life expectancies by age, 1984 Vietnam

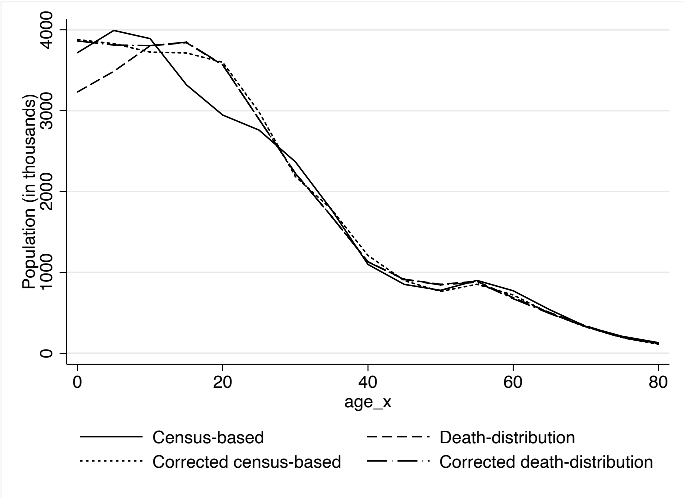

Figure 3. Projected populations by age, 1989 Vietnam

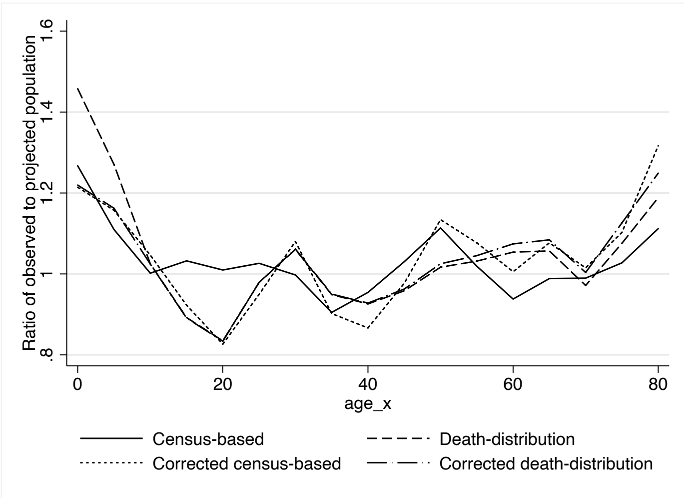

Figure 4. Ratio of observed to projected populations by age, 1989 Vietnam

2.3.1 When to use the iterative method of mortality estimation?

The results show the sensitivity of life tables to three direct mortality estimation methods. In particular, this article advances knowledge of mortality estimation and projection methods by demonstrating how the future growth and distribution of populations may differ under different assumptions derived from the same population inputs.

Compared with the iterative procedure (and in the case of Vietnam), the original census-based method underestimates life expectancy in infancy but also overestimates it after age 45, partly because of its sensitivity to the age structure recorded between the two periods. The iterative variation of the census-based method helps to reduce the influence of the age structure of the population on life expectancy estimates.

When migration and differential census coverage affect the second age distribution, one should use the iterative method to correct growth rates and project the population forward in $t2$. Analysis of the Vietnamese data shows that the use of the iterative procedure results in greater accuracy in life expectancy and projected population, regardless of the completeness of death registration in destabilized populations. For these reasons, `ilt` should be used with the `iterative` option (the default) whenever there is reason to suspect that the input data are deficient or inaccurate, which is often the rule rather than the exception.

3 The ilt command

3.1 Syntax

`ilt, interval(`*numlist*`)` [`census` `death` `iterative` `residual` `life` `proj` `ratio`]

3.2 Description

The most important quantities in life tables are life expectancies, which summarize the average number of years that a cohort l_x will live beyond age x if the mortality conditions observed in a given period remain stable throughout life. `ilt` calculates the functions $(l_x, L_x, T_x, d_x, q_x, e_x)$ of single-decrement life tables following the methods and procedures described in Preston and Bennett (1983), Bennett and Horiuchi (1984), and Merli (1998). The `residual` option calculates the residual number of age-specific migrants. The inputs for `ilt` are

1. the populations aged x to $x+n$ at times $t1$ and $t2$; and
2. the average annual number of deaths between $t1$ and $t2$.

An appropriate dataset must have the same structure and ordering of variables as described in section 2 before they are entered into Stata.

3.3 Options

`interval(`*numlist*`)` lists the length of the interval (in years) between two population distributions. For example, in a situation with 2 censuses 10 years apart, specify `interval(10)`. `interval()` is required.

`census` computes life tables using the census-based method. If this option is specified, `ilt` uses the two population distributions entered and estimates life expectancies starting at the second age group.

`death` uses the death-distribution method to estimate single-decrement life tables.

`iterative` uses the iterative procedure suggested by Merli (1998) to compute the functions of the census-based and death-distribution methods with corrected growth rates. It ensures the convergence between these two complementary estimation techniques. `iterative` is the default method of mortality estimation when none is specified.

`residual` calculates the difference and the ratio between observed and projected—using the corrected death-distribution method—populations at time $t2$. It accounts for the residual number of migrants by age.

`life` produces a comparative graph of life expectancies by age according to the three methods. It must be called with the options `census`, `death`, and `iterative` at the same time. Example: `ilt, interval(10) census death iterative life`.

`proj` produces a comparative graph of projected populations by age according to the three methods. It must be called simultaneously with the options `census`, `death`, and `iterative`. Example: `ilt, interval(10) census death iterative proj`.

`ratio` produces a comparative graph of the ratios between observed and projected populations by age according to the three methods. It must be called simultaneously with the options `census`, `death`, and `iterative`. Example: `ilt, interval(10) census death iterative ratio`.

3.4 Output

As a minimum, `ilt` displays age-specific growth rates (r_x) and the most fundamental functions of single-decrement life tables: person-years (L_x), survivors in the cohort (l_x), and life expectancies (e_x) according to intercensal methods.

The program also calculates the number of net migrants (`residual` option) using the strategy of the United Nations (1970) and generates a set of comparative graphs of life expectancies (`life`), projected populations (`proj`), and ratios between observed and projected populations (`ratio`) according to the three methods (`census, death, iterative`).

3.5 Stored results

Life table functions are stored as Mata colvectors and as Stata matrices under the following names:

Stata matrices

- `Table_1` — life table functions derived from the census-based method
- `Table_2` — life table functions derived from the death-distribution method
- `Table_3` — life table functions derived from the death-distribution method with corrected growth rates
- `Table_4` — life table functions derived from the census-based method with corrected growth rates
- `Table_5` — observed and projected populations—using the corrected death-distribution method—at time $t2$ and by age x and the difference and ratio between them; the difference represents the net number of migrants

Mata vectors

- `r`, `ri` — age-specific growth rates derived in (1) in the census-based method (`r`) and according to the iterative procedure (`ri`)
- `l`, `ld`, `li`, `lc` — individuals in the cohort who are at age x according to the census-based method (`l`), the death-distribution method (`ld`), the death-distribution method with corrected growth rates (`li`), and the census-based method with corrected growth rates (`lc`)
- `L`, `Ld`, `Li`, `Lc` — person-years lived between age x and $x + n$ according to the census-based method (`L`), the death-distribution method (`Ld`), the death-distribution method with corrected growth rates (`Li`), and the census-based method with corrected growth rates (`Lc`)
- `T`, `Td`, `Ti`, `Tc` — person-years lived above age x according to the census-based method (`T`), the death-distribution method (`Td`), the death-distribution method with corrected growth rates (`Ti`), and the census-based method with corrected growth rates (`Tc`)
- `d`, `di` — individuals dying between ages x and $x + n$ according to the death-distribution method (`d`) and the death-distribution method with corrected growth rates (`di`)
- `qd`, `qi` — probability of dying between ages x and $x+n$ according to the death-distribution method (`qd`) and the death-distribution method with corrected growth rates (`qi`)
- `e`, `ed`, `ei`, `ec` — life expectancy at age x according to the census-based method (`e`), the death-distribution method (`ed`), the death-distribution method with corrected growth rates (`ei`), and the census-based method with corrected growth rates (`ec`)
- `pce`, `pcc`, `p`, `pc` — survivorship ratios at age x according to the census-based method (`pce`), the corrected census-based method (`pcc`), the death-distribution method (`p`), and the death-distribution method with corrected growth rates (`pc`)
- `N_projce`, `N_proj`, `N_projcc`, `N_projc` — projected populations at $t2$ and age x according to the census-based method (`N_projce`), the death-distribution method (`N_proj`), the corrected census-based method (`N_projcc`), and the death-distribution method with corrected growth rates (`N_projc`)

Mata scalars

- `n` — age interval length; usually equals five years
- `it_number` — number of iterations required to achieve convergence in growth rates between the census-based and death-distribution methods

4 Conclusions

Table 1 summarizes the methods, data requirements, assumptions, results, and procedures required to compute single-decrement life tables. It overviews each procedure and its purpose and highlights which bibliographic sources provide the formulas used in the manuscript and the conditions under which the program produces valid mortality and migration estimates.

Table 1. Overview of direct methods of mortality estimation generating single-decrement life tables

Method	References	Data required	Assumptions	Main output	Command[a]
Census based	Preston and Bennett (1983)	Two population distributions, organized into five-year age groups	No undercounting, age distortion, or intercensal migration	$r_x(t1, t2)$; l_x; and e_x (starting at $x = 5$)	`ilt, interval(10) census`
Death distribution	Preston et al.'s (1996) version of Bennett and Horiuchi (1984)	Same as the census-based method plus the average annual number of deaths between two periods	Undercount of the population and deaths are constant by age; `ilt` assumes that $L_0 = nl_5 + 0.78d_0$ and $L_{80+} = l_{80} \log_{10}(l_{80})$	$r_x(t1, t2)$, l_x, e_x (at all ages), d_x, and q_x	`ilt, interval(10) death`
Iterative method	Merli (1998)	Same as for the death-distribution method	The initial population is correctly enumerated; convergence is achieved when growth rates in successive iteration cycles are below $1e^{-10}$	Life tables corrected for coverage errors and intercensal migration	`ilt, interval(10) iterative`
Computation of migrants	Merli (1998), United Nations (1970)	Same as for the death-distribution method	Differences btw. obs. and proj. pops. represent migrants[b]	Difference and ratio between obs. and proj. pops. by age	`ilt, interval(10) residual`

NOTES: $r_x(t1, t2)$ stands for age-specific growth rates; l_x are individuals in the cohort surviving to age x; e_x is the life expectancy of individuals at age x; d_x is the number of deaths in the life table; and q_x represents the probability of dying between ages x and $x + n$.

[a] Assuming that populations are 10 years apart.

[b] Rather than differences in completeness of death registration or coverage of observed populations.

The measurement of mortality performed here uses observed data corrected for various sources of error. The `ilt` command produces single decrement life tables describing the mortality conditions of a given population based on only two age distributions (census-based method). This calculation is further extended when the average number of deaths between these distributions is also provided (death-distribution method).

The census-based method is sensitive to errors in the data; while the death-distribution method provides consistent mortality estimates even in the presence of differential census coverage, intercensal migration, and incomplete death registration. `ilt` has the virtue of correcting the sensitivity of the census-based method to differential completeness of census enumeration and residual emigration and of providing corrected growth rates for the death-distribution method (iterative procedure). After iterations, the estimated life expectancies by age converge to similar values in the census-based and death-distribution methods.

However, these methods are not immune to situations in which the completeness of death registration is not constant across ages, nor are they immune to violations of their other assumptions. For example, the iterative procedure assumes that the baseline population is correctly enumerated prior to its forward projection. It also assumes that the age distribution of deaths is centered within the intercensal period. All methods assume that there are no errors in the age reporting of the living and the dead. If this is not the case, the method may produce biased results.

Future developments should investigate the effect of deviations from these assumptions on the sensitivity of mortality estimates. Such studies would need to simulate different sources of error and compare the results with those observed under ideal conditions, similar to what Hill, You, and Choi (2009) and Murray et al. (2010) did. Future enhancements to `ilt` could include, for example, solutions for dealing with differential completeness of deaths by age and the possible inclusion of confidence intervals to capture uncertainties in sampling and measurement.

5 Acknowledgments

I received financial support from the National Council for Scientific and Technological Development, process 304521/ 2019-01. The production of `ilt` benefited from discussions on Statalist and the Stata Blog. These venues remain the most valuable sources of concise and accurate answers to various challenging technical questions. I thank the managing editor, Stephen P. Jenkins, and an anonymous reviewer for their insightful comments. As always, the views expressed in this article are those of the author. Any remaining errors are my sole responsibility.

6 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 23-3
. net install st0729      (to install program files, if available)
. net get st0729          (to install ancillary files, if available)
```

7 References

Baum, C. F. 2006. Stata tip 37: And the last shall be first. *Stata Journal* 6: 588–589. http://doi.org/10.1177/1536867X0600600411.

Bennett, N. G., and S. Horiuchi. 1981. Estimating the completeness of death registration in a closed population. *Population Index* 47: 207–221. https://doi.org/10.2307/2736447.

———. 1984. Mortality estimation from registered deaths in less developed countries. *Demography* 21: 217–233. https://doi.org/10.2307/2061041.

Castanheira, H. C., and J. H. C. Monteiro da Silva. 2022. Examining sex differences in the completeness of Peruvian CRVS data and adult mortality estimates. *Genus* 78(3): 1–25. https://doi.org/10.1186/s41118-021-00151-5.

Hill, K. 2001. Methods for measuring adult mortality in developing countries: A comparative review. Working paper, Johns Hopkins University. http://jhir.library.jhu.edu/handle/1774.2/914.

Hill, K., D. You, and Y. Choi. 2009. Death distribution methods for estimating adult mortality: Sensitivity analysis with simulated data errors. *Demographic Research* 21: 235–254. https://doi.org/10.4054/DemRes.2009.21.9.

Jann, B. 2016. Creating LaTeX documents from within Stata using texdoc. *Stata Journal* 16: 245–263. https://doi.org/10.1177/1536867X1601600201.

Merli, M. G. 1998. Mortality in Vietnam, 1979–1989. *Demography* 35: 345–360. https://doi.org/10.2307/3004042.

Muniz, J. O. 2020. Multistate life tables using Stata. *Stata Journal* 20: 721–745. https://doi.org/10.1177/1536867X20953577.

Murray, C. J., J. K. Rajaratnam, J. Marcus, T. Laakso, and A. D. Lopez. 2010. What can we conclude from death registration? Improved methods for evaluating completeness. *PLOS Medicine* 7: e1000262. https://doi.org/10.1371/journal.pmed.1000262.

Ortega, A. 1982. Tablas de mortalidad. Technical Report 1008, Centro Latinoamericano de Demografía, San José, Costa Rica. https://hdl.handle.net/11362/8718.

Palloni, A., and R. Kominski. 1984. Estimation of adult mortality using forward and backward projections. *Population Studies* 38: 479–493. https://doi.org/10.2307/2174136.

Preston, S. H., and N. G. Bennett. 1983. A census-based method for estimating adult mortality. *Population Studies* 37: 91–104. https://doi.org/10.2307/2174382.

Preston, S. H., and A. J. Coale. 1982. Age structure, growth, attrition, and accession: A new synthesis. *Population Index* 48: 217–259. https://doi.org/10.2307/2735961.

Preston, S. H., I. T. Elo, I. Rosenwaike, and M. Hill. 1996. African-American mortality at older ages: Results of a matching study. *Demography* 33: 193–209. https://doi.org/10.2307/2061872.

Preston, S. H., P. Heuveline, and M. Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell.

Queiroz, B. L., M. R. Gonzaga, A. M. N. Vasconcelos, B. T. Lopes, and D. M. X. Abreu. 2020. Comparative analysis of completeness of death registration, adult mortality and life expectancy at birth in Brazil at the subnational level. *Population Health Metrics* 18 (Suppl. 1)(11): 1–15. https://doi.org/10.1186/s12963-020-00213-4.

Spoorenberg, T. 2008. What can we learn from indirect estimations on mortality in Mongolia, 1969–1989? *Demographic Research* 18: 285–310. https://doi.org/10.4054/DemRes.2008.18.10.

United Nations. 1956. Manual III: Methods for population projections by sex and age. ST/SOA/Series A, Population Studies 25, United Nations—Department of Economic and Social Affairs, New York.

———. 1970. Manual VI: Methods of measuring internal migration. ST/SOA/Series A, Population Studies 47, United Nations—Department of Economic and Social Affairs, New York.

About the author

Jerônimo O. Muniz is an associate professor and researcher in the Department of Sociology at the Federal University of Minas Gerais, Brazil. His research interests include formal demography, forecasting methods, social stratification, and sociology of race. He is also the author of `mslt`, a command that computes the functions of multistate life tables.