SSC Annual Meeting, June 2001
Proceedings of the Survey Methods Section, pp. 99-102

WHAT'S NEW FOR THE 1995-1997 CANADA LIFE TABLES

Brad Thomas, Doreen Duchesne, Marek Wysocki and Robert Bourbeau[1]

ABSTRACT

A life table shows expected mortality experience for a hypothetical cohort of 100,000 persons born at the same time and subjected to the age-sex-specific mortality rates which an actual population experienced. The last set of detailed life tables for Canada and the provinces was for the 1990-92 period (Statistics Canada Cat. No. 84-537), based on the 1991 Census of Population (CEPOP) and mortality data for 1990 through 1992. The availability of 1996 CEPOP and 1997 mortality data now allows us to produce detailed life tables for the 1995-97 period. The previously cited Statistics Canada publication describes in detail the methodology of the 1990-92 life tables, to which we made a number of improvements for the 1995-97 life tables. The focus of this paper will be on the most significant improvements, namely the use of a model for estimating probabilities of dying at ages over 87 (Coale and Kisker, 1990).

KEY WORDS: Life table; Mortality.

RÉSUMÉ

Une table de mortalité du moment est un modèle décrivant l'extinction progressive d'une cohorte fictive de 100 000 personnes, sous l'effet des taux de mortalité par âge et par sexe calculés au cours d'une période donnée. Les dernières tables de mortalité complètes produites pour le Canada et les provinces couvraient la période 1990-1992 (Statistique Canada numéro 84-537 au cat.). Ces tables étaient basées sur les décès observés au cours de la période triennale 1990-1992 et sur les effectifs de la population recensée en 1991 (REPOP de 1991). La disponibilité des données du recensement de 1996 (REPOP de 1996) et des décès de 1995 à 1997 nous permet d'établir les tables de mortalité complètes pour la période 1995-1997. La publication de Statistique Canada citée précédemment fournit une description détaillée de la méthodologie utilisée pour bâtir les tables de 1990-1992; nous avons apporté plusieurs modifications à ces dernières dans l'établissement des tables de 1995-1997. Cet article portera sur les améliorations les plus importantes, notamment l'utilisation d'un modèle pour estimer la probabilité de décès au-delà de 87 ans (Coale et Kisker, 1990).

MOTS CLÉS: Mortalité; tables de mortalité.

1.0 INTRODUCTION

A life table represents a universally accepted demographic or actuarial model that synthesizes the mortality experience of a population, in a clear and concise manner, and enables comparative measures of expected longevity. In the construction of these tables, it is customary to assume that a hypothetical cohort of 100,000 individuals born at the same moment in time is subject to the age-sex-specific mortality rates experienced by an actual population during a specific time period. In Canada, these life tables are published every 5 years, and they cover a period of 3 years, centred around a year in which a Census of Population was taken. Although there was a Census of Population this year (2001), the most recent life tables centre around the previous Census of Population, which was in 1996.

This article will highlight the most significant changes between the methods used to construct the 1990-1992 Canada Life Tables and the 1995-1997 Canada Life Tables. Details on the methodologies used in each publication are found in Statistics Canada (1995) and Statistics Canada (2001b).

2.0 CHANGES IN METHODOLOGY FROM 1990-92 TABLES TO 1995-97 TABLES

The quantity of greatest interest in the life table is the conditional probability:

$$Q(x) = \Pr\{\text{ Individual dies before age } x+1 \mid \text{Individual has lived to age } x\ \}.$$

The life table shows values of $Q(0)$ through a final value which varies between $Q(100)$ and $Q(109)$, depending on the province and sex combination.

[1]Brad Thomas, Doreen Duchesne and Marek Wysocki, Household Surveys Methods Division, Statistics Canada, R.H. Coats Building, 16th floor, Tunney's Pasture, Ottawa, Ontario, Canada, K1A 0T6, and Robert Bourbeau, Professeur agrégé, Département de démographie, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec, Canada, H3C 3J7, thombra@statcan.ca.

The methodology followed in constructing the 1995-97 tables is the same as that employed previously in producing the set of tables for the years 1990-92, but with two major changes. These occur at the extreme ends of the age range in the complete life tables: (1) we have used a model for the very old ages (88 and above), and (2) we have used January 1 population counts instead of birth counts for the very young ages (0 through 4). There are some lesser modifications as well, but these changes are the focus of this article.

2.1 Methodology for the Very Old Ages (over 87)

It is generally accepted that population estimates become very uncertain at high ages. Coale and Kisker (1990) found significant problems in accurate estimation of populations for high ages in the U.S. in part due to misreporting of age and so they developed a model-based alternative for estimating the central death rates for advanced ages. This method also gives a more realistic evolution of mortality rates at very old ages, since recent observations show a deceleration in the rate of increase of mortality rate with increasing age. This method follows these steps:

Step 1

Within each province and sex combination, compute age-specific mortality rates for age 82 <= x <= 86

$$M(x) = \frac{1}{P_{1996}(x)} \left(\frac{D_{1995}(x) + D_{1996}(x) + D_{1997}(x)}{3} \right)$$

where $P_{1996}(x)$ is the 1996 Census of Population count (mid-year) at age x for that province and sex combination, and $D_y(x)$ is the death count for year y of persons aged x for that province and sex combination. Note that rather than take the death counts just for 1996, we use a 3-year average over 1995 - 1997 for greater stability.

Step 2

Take average over the age range 82 <= x <= 86

$$\bar{M}_{84} = \frac{1}{5}(M(82) + M(83) + M(84) + M(85) + M(86))$$

Step 3

Define constant

$$K_{85} = \frac{1}{4} \log\left(\frac{M(86)}{M(82)} \right).$$

Step 4

Define another constant

$$S = -\frac{1}{325} \left(\log\left(\frac{\bar{M}_{84}}{M(110)} \right) + 26\ K_{85} \right).$$

Furthermore, we choose to impose $M(110)$ = 1 for males, 0.8 for females (as documented by Coale and Kisker (1990) in one of their applications).

Step 5

For 87 <= x <= 117, make new CK (Coale-Kisker) estimates for crude death rate, using the constants defined in **Step 3** and **Step 4**:

$$M_{CK}(x) = \bar{M}_{84}\ EXP\left[K_{85}\ (x - 84) + (x - 84)\ (x - 85)\ \frac{S}{2} \right]$$

Step 6

For 87 <= x <= 117 compute new $Q(x)$:

$$Q_{CK}(x) = \frac{2\ M_{CK}(x)}{(2 + M_{CK}(x))}.$$

This left us with the question of when to replace the $Q(x)$ calculated from reported data with this new model-based $Q_{CK}(x)$. For the province-level estimates, we chose to keep $Q(x)$ up to age $x = 87$, and then use $Q_{CK}(x)$ starting with $x = 88$. At the national level, however, we had confidence to keep the data-based $Q(x)$ to $x = 92$, and then replace with $Q_{CK}(x)$ starting at $x = 93$. At the national level, the model-based estimates had the desired effect for males, namely $Q(x) < Q_{CK}(x)$. Thus if there was exagerration of age among males, then $Q(x)$ would be too low and $Q_{CK}(x)$ corrects that. The model-based estimates actually had the opposite effect for females. Nonetheless, we felt more comfortable using the model for its stability at the advanced ages for both sexes.

2.2 Methodology for the Very Young Ages (0 - 4)

As described in both Statistics Canada (1995) and Statistics Canada (2001b), the basic formula for $Q(x)$ for $0 <= x <= 4$ in both the 1990-92 and 1995-97 life tables is:

$$Q(x) = 1 - (P'(x) / E(x)) (E(x+1) / P''(x))$$

where:

- $P'(x)$ is the number of persons who attained age x during the three-year mortality period of 1995-97 and who were alive at the end of the year in which exact age x was attained
- $E(x)$ is the number of persons who attained age x during the 1995-97 period
- $P''(x)$ is the number of persons alive at the end of the calendar year in which age x was attained and whose $(x+1)$th birthday falls in the 1995-97 period

Each of the P', E and P'' quantities are first computed for individual years and then summed up:

$$E(x) = E^{1995}(x) + E^{1996}(x) + E^{1997}(x)$$

where

- $E^z(x)$ is the number of persons who attained age x in year z

and so on.

Now let

$$P^z(x) = (\text{\# persons of age } x \text{ at the start of year } z)$$

For the 1990-92 tables, the methodology was to start with birth counts for year z, which would be $E^z(0)$, and then use death counts to model how this population would change over time to get $P^{z+1}(0)$, $P^{z+2}(1)$, and so on. But the model did not take immigration into account, and so it produced a series such that the following relationship was forced to occur:

$$E^z(0) >= P^{z+1}(0) >= P^{z+2}(1) >= P^{z+3}(2) >= \ldots \quad (*)$$

But the Statistics Canada 1 January population estimates account for immigration (e.g. Statistics Canada (2001a)), and they did not uphold the relationship (*) in a number of cases. In addition the death counts used to obtain the P-series from $E^z(0)$ contain deaths among immigrants.

For the 1995-97 tables, we chose instead to use the Statistics Canada 1 January population estimates directly for the $P^z(x)$, and then model the $E^z(x)$ as needed. As a result, the relationship (*) was not forced to occur, and this was a better reflection of reality.

3.0 FOR LIFE TABLES 2000-02

Some further improvements are already being considered for the next set of life tables, which will centre around the 2001 Census of Population, and cover the period of 2000-02.

3.1 More Sophisticated Graduation Techniques

The changing of methods for estimating $Q(x)$ between different age ranges can result in some harsh jumps where one method ends and another one takes over. In addition, although one anticipates differences among the provincial $Q(x)$ series, as well as between the national level $Q(x)$ and those at the provincial level, the "smoothness" obtained at the national level should ideally be carried over to the provincial level. There are sophisticated graduation techniques available in actuarial literature (such as the Wittaker-Henderson Type B difference-equation formulae in London (1985)) which should be explored.

3.2 Variance Estimation

Once we obtained our estimate of Q(x), we calculated an estimate of Var(Q(x)) following the method given in Chiang(1984), namely:

$$var(Q(x)) = \frac{1}{\bar{D}(x)} Q(x)^2 (1 - Q(x))$$

where we use the average death count over the 1995-1997 period:

$$\bar{D}(x) = \frac{1}{3} (D_{1995}(x) + D_{1996}(x) + D_{1997}(x))$$

as we did back in Section 2.1. There is zero covariance between the $Q(x)$ for different ages (x-values) within a province and sex combination where the $Q(x)$ have been estimated using a census rather than a sample of death certificates, etc..

Chiang derives this variance estimation and the zero covariance property for estimation of
the $Q(x)$ from death population counts for just that age value, x. For example, if you wish to estimate $Q(87)$ for males, then you would use just the death count and

population count data for age 87 for males.

In our case, however, we used different smooting techniques to estimate the $Q(x)$, such as using death and population counts from neighbouring ages. For example, in order to estimate $Q(87)$ for males, we used not only death and population count data for males age 87, but also death and population count data for males age 82 and 92. This makes it harder to justify the zero covariance assumption, because even simple linear combinations of statistically independent random variables may not preserve statistical independence. (E.g.: if $X1$ and $X2$ are statistically independent, then $Y1 = X1 + X2$ and $Y2 = X1 - X2$ will not be statistically independent.) So our estimates of $Q(82)$ and $Q(87)$ are not statsitically independent of each other. Further study into the variance properties of our $Q(x)$ estimates would be worthwhile.

ACKNOWLEDGEMENTS

The authors would like to thank L.M. Cornelis of the Office of the Chief Actuary of Canada for his suggestion for the Whittaker Henderson Type-B difference-equation formulae during his review of the manuscript for Statistics Canada (2001b).

The authors would also like to thank Pierre David of the Social Survey Methods Division of Statistics Canada and François Nault of the Demography Division of Statistics Canada for their helpful comments on earlier drafts of this paper.

REFERENCES

Chiang, C.L. (1984). *The Life Table & its Applications.* Malabar, Florida: RE Krieger Pub. Co.

Coale, A.J. and E.E. Kisker (1990). «Defects in data on old-age mortality in the United States: New procedures for calculating mortality schedules and life tables at the highest ages». *Asian and Pacific Population Forum* 4(1):1-31.

London, Dick (1985). *Graduation: the Revision of Estimates.* Winsted, Conneticut: ACTEX Pub.

Statistics Canada (1995). *Life Tables, Canada and Provinces, 1990-92,* Cat. No. 84-537. Ottawa, Ontario: Minister of Industry.

Statistics Canada (2001a). *Annual Demographic Statistics, 2000,* Cat. No. 91-213. Ottawa, Ontario: Minister of Industry.

Statistics Canada (2001b). *Life Tables, Canada, Provinces and Territories, 1995-97,* Cat. No. 84-537. Ottawa, Ontario: Minister of Industry.