Introduction

Guillaume Wunsch

"To be or not to be?" mankind has probably raised this question since the first men and women confronted life and death on earth. Will Saturn bring me old age? Though the answer depends upon the gods, the evil spirits, or disease, according to the times, Man has nevertheless sought his future in the leaves of the tea-cup, the palm of one's hand, the crystal ball, or ... the life table. The history of the life table has been briefly sketched by D. Smith and N. Keyfitz (1977). Though the origins of the "mortality table" (as the French say) date back to the classic studies of Graunt, Halley, and Euler, a third century A.D. table of annuities, attributed to Ulpian, bears witness to the interest of the Romans for life annuities and therefore for compiling life experiences. Indeed, life tables are a subject of interest not only for demographers but also for actuaries and epidemiologists. The study of the extinction of a group of "lives" forms an important domain of insurance theory, and the construction of the life table is described in all actuarial books dealing with life insurance; for a recent example, see F.E. De Vylder (1997). Even if nowadays non-life insurance problems dominate actuarial theory, life contingencies still form the backbone of the insurance business. Life tables are also considered in epidemiology; see e.g. the textbook by J. Estève et al. (1993). Epidemiologists are however more interested in measuring the incidence and prevalence of diseases, and determining possible risk factors of morbidity and mortality, than in evaluating the mortality of the general population. As a corollary, epidemiology draws its data more from special surveys and registers, such as longitudinal heart studies or cancer registries, than from general population statistics such as vital registration and census.

Demographers are interested in the life table because the latter summarises the mortality experience of a birth cohort or of a period, mortality being with natality and migration one of the three factors influencing population change. Life tables

G. Wunsch et al. (eds), The Life Table, Modelling Survival and Death, 1-12

are also the basis for stationary and stable population models. Most life tables relate to a specific period, in a cross-sectional approach, and the life table in this case is thus based on the risks of dying by age and sex during the period. All other indicators, such as the life expectancy at a specific age, are entirely derived from these risks. A period expectation of life at birth can therefore be considered as a summary measure of all the risks of dying by age. It is thus a useful indicator of the mortality level of a given period, controlling for the age structure of the population. As such, the period mean length of life, in the fictitious cohort approach, can be used instead or in addition to a standardised mortality rate. Both measures indicate the mortality level of a population independently from its population pyramid, contrary to the crude death rate that is highly influenced by the demographic age structure.

In demography, the life table has often been used for descriptive purposes, as a tool for deriving the expectation of life at birth for a specific period, in order to compare mortality levels over time or space. For instance, the table constructed by Edmund Halley "exhibits the number of people in the City of Breslaw of all ages, from the birth to extream old age, and thereby shews the chances of mortality at all ages", though Halley also proposed using the life table "to make a certain estimate of the value of the annuities for lives" (E. Halley, 1693). The Breslau table therefore also represented for Halley some sort of standard or model of mortality, which could be applied to other populations at other places. In epidemiology, on the other hand, the life table is used to estimate survival times for individuals in a particular group, and to compare the survival data between e.g. treatment and control groups. The emphasis in this case is laid more on the impact of censoring and on hypothesis testing than on the comparison of mean lengths of life, in view of obtaining some generalizable conclusion on the effect of the treatment. In this respect, the objectives of epidemiology are closer to those of actuarial studies, as stipulated above by Halley, than to the descriptive case-study approach usually adopted by demographers.

More specifically, as Nathan Keyfitz (1968) points out, the life table is a population model covering the simple case of a group of people (or birth cohort) born at the same moment, closed to migration, and followed through successive ages until they die. Keyfitz also stresses the fact that "the life table is a scheme for expressing the facts of mortality in terms of probabilities" (Keyfitz, *op. cit.*). One could add that the life table furthermore assumes the homogeneity of groups, i.e. that all subjects have the same distribution of survival times

(S. Anderson et al. 1980). The model is also non-parametric or distribution-free, in the sense that it does not require specific assumptions to be made about the underlying distribution of the survival times (D. Collett, 1994). It is worth considering briefly some other characteristics of this elementary model.

As one knows, cohort analysis as exemplified by the life table, is a special form of longitudinal analysis where data are structured according to one's period of birth under the assumption, as N.B. Ryder (1965) has stressed, that members of a birth cohort share "a common historical location" as they have lived through similar experiences. Thus, in Ryder's terms, "each cohort has a distinctive composition and character reflecting the circumstances of its unique origination and history; (...) the community of date equips each cohort with its own expanse of time, its own style, and its own truth". For example, wars or epidemics may leave a negative imprint on the future health of individuals who have experienced these hardships at a similar age, or unfavourable health conditions early in life might leave the strongest alive (thus violating the homogeneity assumption) and lead to reduced cohort mortality later in life. These and other possible cohort effects are discussed in J. Hobcraft et al. (1982). Actually, the choice between a cohort or a period approach depends on the phenomenon one is studying and on the questions raised; cohort effects seem to be less common for example in the field of fertility, where period effects predominate (M. Ni Bhrolchain, 1993).

Moreover, consider the risk of dying at a certain age. This risk is age-dependent, as mortality varies greatly according to age-related biological factors. It is also period-dependent, as the conditions of the period, in the social, economic, and public health fields for instance, have an impact on mortality. It is finally cohort-dependent, as the similar history of the members of the cohorts in such matters as vaccination, availability of medicinal drugs, smoking and eating habits, etc., have an impact on their eventual mortality. One can therefore consider that the risk of dying is subjected to age effects, period effects, and cohort effects, and that the age-cohort approach as stressed in the life table model, or its period equivalent if a synthetic cohort approach is used, does not encompass the true nature of the underlying phenomena. *Age-period-cohort modelling* would in this case be a better alternative to the life table, either cohort or period, if one is interested in understanding the nature of mortality changes over time.

In addition, cohort analysis as subsumed by Keyfitz' definition of the life table model, refers to what is called a *repeated cross-sectional design*, in the sense that the experience of a group of persons – and not individual life trajectories – is

followed through time. This means losing the valuable information on individual differences *within* the cohort, that is on intra-cohort heterogeneity. If individual longitudinal data are available, it is thus strongly recommended to use the information on individual life histories, instead of aggregating the data by cohort as in the life table model, and to apply *event history analysis* (such as the Cox proportional hazards model) to the data on individual survival times. Cohorts can still be distinguished in this approach by including in the statistical model the year of birth (or the calendar period defining the cohort) among the individual explanatory variables, similarly to such categorical variables as race or place of birth. The converse is obviously not possible: if one has only aggregate data, it is impossible to break the latter down into individual life courses.

The main advantage of a longitudinal approach, whatever the method, is to locate events in time (see e.g. H.P. Blossfeld and G. Rohwer, 1995). The temporal sequence of events can indeed indicate possible causal patterns, as causes always have to precede effects in time. In the field of longitudinal mortality research, two of the present editors (G. Wunsch and J. Duchêne, et al., 1996) have suggested that there are advantages in taking account not only of the order of events but also of the type of states one has experienced during one's lifetime, a state being defined by the prevalence of the various characteristics of an individual at a point in time, and finally the time spent in each state. Such a *state-time-order transition model* requires however individual longitudinal data, which once again the life table does not. The life table remains therefore a useful tool if only aggregate data are available.

The *homogeneity* assumption is not very realistic in practice. Life tables are therefore often constructed for sub-groups, which are presumed to be more homogeneous as to the risks of dying. Breaking down the population into more homogeneous sub-groups leads to diminishing numbers of people at risk and therefore to increasing confidence intervals for the values of the survivor functions. Moreover, one cannot control all possible confounding factors, for the simple reason that some are unknown to us. Unobserved heterogeneity is therefore always present in the data; for this reason, randomised trials are much to be preferred, as the unknown causal factors are much less likely to wreak havoc in this case on the results (J.M. Elwood, 1988). Randomised trials are however expensive to conduct and often difficult to use in practice, among others for ethical reasons. They are rarely, if ever, set up in demography but are recognized as "the epitome of scientific validity" (B. Andersen, 1990) in medical research.

Life tables have traditionally been computed according to age and sex. Age is the duration of one's life, and furthermore risks of dying vary considerably according to age. Mortality varies also according to gender, thereby justifying its inclusion in life table analysis. Age and sex are furthermore always available in vital registration and census data. Mortality varies however according to many factors other than age and gender. Period life tables have therefore been calculated by regions, socio-economic categories, ethnic groups, etc. Differential mortality studies have shown that many variables are indeed associated with mortality; these studies are unfortunately limited by the type of data available. Death certificates include very few variables other than age at death and sex. Differential mortality studies are therefore usually based on some form of record linkage between vital registration and the census, the latter including many more individual variables than the vital registration system. In some cases, such as in the study of Norwegian mortality by G. Wunsch et al. (1996), record linkage can even be performed between successive censuses, yielding information on changes in the characteristics of individuals over their lifetime.

As mortality is often the outcome of a disease process, life tables are also computed by cause of death. In this case, one can take into account the underlying, immediate, or associated causes of death, depending on their place in the disease process. A person might also die from multiple causes. This often happens at high ages where people suffer from several diseases at the same time. In this case, ascribing death to a sole cause does not make much sense. Reducing risks of dying by cause(s) leads to increasing life expectancies. Does an increase in life expectancy also mean that people are healthier than before? In other words, do improvements in life expectancy lead to improving the health of the population? For some, adhering to the expansion-of-morbidity hypothesis, mortality reductions will produce higher expectations of life but also lead to more years with morbidity. For others such as Fries, partisans of the compression-of-morbidity hypothesis and a fixed life span, chronic diseases can be pushed on the contrary towards older ages as mortality falls. A third hypothesis, due to Manton, stresses the fact that increasing survival will yield more years with morbidity but that years with severe morbidity will remain relatively constant as the rate of progression of chronic diseases is reduced. These hypotheses are examined in a recent book by W. Nusselder (1998), who concludes that it is "the subtle interplay of mortality and morbidity that turns the scale in favour of compression or, on the contrary, expansion of morbidity". Changes in mortality should therefore never be examined independently from changes in morbidity.