

Chapter 25

Mortality Modeling Perspectives

Hoang Pham

Department of Industrial and Systems Engineering,
Rutgers University,
96 Frelinghuysen Road,
Piscataway, NJ, 08854, USA

25.1 Introduction

As the human lifespan increases, more and more people are becoming interested in mortality rates at higher ages. Since 1909, the birth rate in the United States has been decreasing except for a major significant increase after World War II, between the years 1946 and 1964 [1], also known as the baby boom period. People born during the baby boom are now between the ages of 44 and 62. According to the National Center for Health Statistics, US Department Health and Human Services, in 1900–1902 [2, 3], one could expect to live for 49 years on average.

Today, an infant can expect to live about 77 years. As of recent years and in prediction, the life expectancy for an infant born may be even higher. With the human lifespan increasing and a large part of the United States population aging, many researchers in various fields have recently become interested in studying quantitative models of mortality rates [4]. Scientists in biological fields are not only interested in organisms and how they are made, they are also interested in what happens to organisms over time. A study of yeast, which would interest biologists, showed the effects of senescence as well as a model that accurately represents the experimental data. It has been shown that the addition of a *Sir2* gene can prolong life in yeast [5]. Once we can model human aging, we can look for ways to extend our lifespan and counteract the negative aspects of aging.

Researchers in the medical field are interested in the cost effectiveness of new medication with respect to old medication. The cost per life year saved can be evaluated for each of the two medications with respect to the mortality model. For this type of work, it is important to use a model that will fit the mortality data for the population in which the medication is to be distributed [6]. This implies that a mortality model be chosen that accurately describes the data in question and indeed very crucial, but that also depends on the modeling decision criteria [7].

In the mathematical and physical fields, researchers are interested in predicting the failure of machines, or other man-made objects [8]. Through a series of experiments and mathematical calculations, one can predict when an object will fail

without a direct test on the object. This is useful because one can use available data to create a model that can predict the remaining lifetime of humans. A direct test would take the entire lifespan instead of the shorter time needed to create and use the model. Physicists and mathematicians are also interested in creating models that will depict how aging and environmental factors will effect the survival of the human population [9].

Social scientists in the United States are currently very interested in projections of life expectancy and aging due to the deteriorating system of social security. The current “Pay-As-You-Go” system, which provides benefits for the elderly, is breaking down as the life expectancy is increasing and a large part of the population is aging [10]. Many organizations, including insurance companies, today rely on mortality models and projections of human aging.

This chapter first introduces the motives as well as the reasons why studying mortality data is important. Some literature reviews on the mortality modeling and analysis are then discussed. Several distributions applicable and common used to human mortality studies are also mentioned.

25.2 Literature Discussions

Gavrilov and Gavrilova [8] discussed a brief overview of reliability theory, relating it to mechanical systems and biological systems. They stated that organisms tend to die according to the Gompertz law, while technical devices tend to die according to the Weibull law. They also brought up the idea that the individuality and uniqueness in humans is caused by different combinations of defects in the organism. Aging is a direct result of system redundancy in organisms, and initial flaws in organisms cause the tendency to follow the Gompertz law. Gavrilov and Gavrilova believed that when the Gompertz law and Weibull law both fail, then mortality is following a more general law in reliability theory. They believed that reliability theory provides a great predictive and explanatory power to theories of aging. Gavrilov and Gavrilova provided insight into several different distributions found in reliability theory and how they relate to organisms as well as machines, but did not compare the distributions against experimental data.

Thatcher [11] explored a model showing that the probability of dying increases with age, and believed that it is valid for both modern and historic data. First, he studied three existing models, the “fixed frailty” model, the stochastic process model, and a theory based on genetics, and tackled the question of why mortality seems to follow the logistic model by analyzing four models such as the logistic model, Gompertz model, Weibull model, and the law of mortality, for effectiveness of modeling mortality at high ages. Out of the four methods, Thatcher found that the logistic model out performed the other three models for data from 1980 to 1982, from England and Wales. He also found that the logistic model fit best for the historic data from these areas. In addition, he described several theories for predicting the probability distribution of the highest age. One view was that there is a fixed

upper limit to the length of human life. The other is that there will continue to be a probability distribution for the highest age given the circumstances. Thatcher used the data from several different areas of the United Kingdom, such as England and Wales, but not from the United States. Since he was studying historic data, he chose an area where he would have access to mortality records from the past, such as the tenth and twelfth centuries. To analyze the models and fit them to the data, he used a modified version of the maximum likelihood estimation method to accommodate for using data from a life table.

Pletcher and Neuhauser [4] studied the creation of criteria for modeling aging, which they felt would create more consistency in the research of aging in different fields. They found five experimental results in the field of biodemography, which impacts the study of aging. They noted that in reliability theory, many models are useful in looking at biological aging, but felt that the models lack biological realism. They proceeded to examine models in the four research areas: molecular biology, physics, reliability engineering, and evolutionary biology/population genetics. They noted that the experimental results from one field of study may fail to model the same data in using another field's criterion. They also outlined a simple mechanistic model that built from all four fields and, through simulations, upheld the experimentally found results. They realized that with further investigation and simulations, their model would not hold up; however, their purpose was to open communication and research between the different fields. Pletcher and Neuhauser realized the need to find a way to join different scientific fields under common criteria.

Bongaarts [12, 13] used the data from the Human Mortality Database for females and males aged 25–109 in 14 different countries, to test the fit of logistic models for the force of mortality. He also proposed a new shifting logistic model that he hoped would better predict age-specific rates of adult mortality. He compared his model to the Lee–Carter method [14, 15] for modeling and forecasting mortality by age. Bongaarts found that his model addressed several weaknesses in the Lee–Carter model, and believed that it provided a basis for age-specific mortality projections. Higgins [16] examined several mathematical models including the Gompertz, Perks, Polynomial, and Wittstein models used to describe and explain human mortality. He discussed the importance of cohorts in models as well as requirements for the ideal model; however, it required a complicated function and is not yet defined.

Pletcher [17, 18] stated that “well-defined statistical techniques for quantifying patterns of mortality within a cohort and identifying differences in age-specific mortality among cohorts are needed”. He examined ways to find the parameters for each of the models using the maximum likelihood estimates. He also set a minimum for the number of individuals needed in a specific cohort to make the experiments effective, and noted that no fewer than 100–500 individuals are needed in each cohort, and that combining cohorts to achieve these numbers are much more helpful in accurately estimating mortalities for age groups, rather than keeping them separate. Pletcher explained how Gompertz parameters were traditionally estimated with linear regression, which has a much higher bias than the maximum likelihood method. He concluded that the extended use of the maximum likelihood methods provide for much better analysis of the mathematical models.

25.3 Mortality Modeling

This section briefly discusses six common used distributions in the area of aging and mortality modeling. Table 25.1 presents a list of common distributions such as Gompertz, Gompertz–Makeham, Logistic, log logistic, loglog and Weibull that commonly used on mortality modeling and analysis. The probability density function (pdf) and hazard rate of each distribution are listed in Table 25.1, column 2 and 3 respectively. The hazard rate function is also known as the failure rate in reliability engineering; the force of mortality in demography, the intensity function in stochastic processes, and the age-specific failure rate in epidemiology.

In the context of reliability modeling, the interrelationships between the pdf $f(t)$, failure rate function $h(t)$, cumulative failure rate function $H(t)$, and reliability function $R(t)$, for a continuous lifetime t can be summarized as

$$\begin{aligned} h(t) &= \frac{f(t)}{R(t)} \\ H(t) &= \int_0^t h(x) \, dx \\ R(t) &= e^{-H(t)} \end{aligned}$$

Note that the cumulative failure rate functions must be a non-decreasing function for all $t \geq 0$ and $\lim_{t \rightarrow \infty} H(t) = \infty$.

The characteristic of the hazard rate function for the occurrence of a particular event can be increased, decreased, constant, bathtub shaped [21], and Vtub shaped [20] that can help to describe the failure mechanism or symptoms. Distribution models with decreasing hazard rate functions are much less common but can find in the applications of learning behaviors. Models with increasing hazard rates often can be used in applications that reflect the aging or wear and tear. The bathtub-

Table 25.1 Probability density function and hazard rate of distributions

Distribution	pdf	Hazard rate
[19]	$f(t) = \beta e^{\left[\theta t - \frac{\beta}{\theta} (e^{\theta t} - 1)\right]}$	$h(t) = \beta e^{\theta t}$
Loglog [20]	$f(t) = \beta \ln(\theta) t^{\beta-1} \theta^{t\beta} e^{1-\theta^{t\beta}}$	$h(t) = \beta \ln(\theta) t^{\beta-1} \theta^{t\beta}$
[21]	$f(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1} e^{-\left(\frac{t}{\theta}\right)^\beta}$	$h(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}$
Gompertz–Makeham	$f(t) = (\gamma + \beta e^{\theta t}) e^{\left[-\frac{\beta}{\theta} (e^{\theta t} - 1) - \gamma t\right]}$	$h(t) = \gamma + \beta e^{\theta t}$
Logistic	$f(t) = \beta e^{\theta t} \left[1 + \frac{\gamma\beta}{\theta} (e^{\theta t} - 1)\right]^{-\frac{\gamma+1}{\gamma}}$	$h(t) = \frac{\beta e^{\theta t}}{\left[1 + \frac{\gamma\beta}{\theta} (e^{\theta t} - 1)\right]}$
Log-logistic	$f(t) = \frac{\beta \theta t^{\beta-1}}{(1 + \theta t^\beta)^2}$	$h(t) = \frac{\beta \theta t^{\beta-1}}{(1 + \theta t^\beta)}$

shaped hazard rate function plays an important role in reliability applications such as human life and electronic devices [22].

In the Gompertz model, the two parameters β and θ are positive; β varies with the level of mortality and θ measures the rate of increase in mortality with age. The Gompertz model is used widely with biologists and demographers [18], and assumes that the mortality increases exponentially with age. This model was developed by Benjamin Gompertz and published in the *Philosophical Transactions of the Royal Society of London* in 1825 [19]. Although this distribution was developed almost 200 years ago, it is still commonly used in modeling biological mortality. In Messori’s study [6] of cost-effectiveness analysis, he states that the Gompertz function best models the mortality data he used from England.

The Gompertz model has been extended with the addition of a constant γ to take into account the background mortality due to causes unrelated to age (see Table 25.1). This model is known as the Gompertz–Makeham model. The model likely represents an improvement over the Gompertz model at younger ages, but it still seems to over-estimate mortality at the oldest ages [11]. The Gompertz–Makeham mortality law is also a widely used distribution in gerontological investigation [23]. In Yu *et al.*’s conclusion, they do not recommend the model for human populations because the distribution needs to be corrected for sex differences.

The Weibull distribution listed in Table 25.1 is a two-parameter distribution; one parameter controls the scale and the other controls the shape of the distribution. Gavrilov and Gavrilova [8] indicate that a difference between the applications of the Weibull and Gompertz distributions is such that technical devices follow the Weibull distribution, while organisms follow the Gompertz. Even though Weibull may have been earmarked for technical devices, due to its flexibility depending on the parameters, it can also be applied to mortality and human population growth modeling. A recent study by Pham and Lai [22] on the generalization of the Weibull can also be applied in modeling the mortality and human population aspects. The Gompertz and Weibull distributions are both commonly used in reliability engineering.

Table 25.2 Mortality modeling and distribution functions and analysis literature

Group	References
General modeling and data analysis	[8, 9, 15, 24–37]
Age-specific mortality	[12, 13, 15, 16, 18, 38–41]
Gompertz and Weibull models	[6, 19, 21, 42–44]
Logistic, Loglog, and other common models	[7, 45]
Measurements, modeling fitting and criteria	[7, 18, 43, 44]
Aging	[4, 5, 8–10, 35, 42, 46–50]
Life expectancy	[13, 15, 29, 43, 44, 51]
Mortality modeling	[15–17, 19, 24, 34, 36–40, 43, 44, 47, 48, 52–59]
Population data	[6, 14, 38, 41, 60]
Forecasting	[24, 43, 60, 61]

The three-parameter logistic model is derived from the work of Pierre Verhulst in the early 1800s and is commonly used in many disciplines including mortality human population, software reliability engineering, human factor analysis, cancer medicine applications and applied engineering statistics. In Thatcher's analysis of highest attained age [11], he noted how the logistic model better predicted higher ages for his data. In including this model, this study will be able to show whether or not this model can also accurately predict for the entire age range, not just the highest ages. A summary of references of research papers and books on mortality modeling and analysis is given in Table 25.2.

References

1. United States Center for Disease Control and Prevention. Live births, birth rates, and fertility rates, by race: United States. 1909–2000. 2000.
<http://www.cdc.gov/nchs/data/statab/t001x01.pdf>
2. United States National Center for Health Statistics
3. US Census Bureau, Population Division. <http://www.census.gov/population/cen2000/phc-t3/tab03.pdf>
4. Pletcher SD, Neuhauser C (2000) Biological aging – criteria for modeling and a new mechanistic model. *International Journal of Modern Physics C* 11:525–546
5. Gillespie CS *et al.* (2004) A mathematical model of ageing in yeast. *Journal of Theoretical Biology* 229:189–196
6. Messori A (1997) Survival curve fitting using the Gompertz function: a methodology for conducting cost-effectiveness analysis on mortality data. *Computer Methods and Programs in Biomedicine* 53:175–164
7. Pham H, Deng C (2003) A predictive-ratio risk criterion for selecting software reliability models. The 9th ISSAT Proceedings on Reliability and Quality in Design, August
8. Gavrilov LA, Gavrilova NS (2001) The reliability theory of aging and longevity. *Journal of Theoretical Biology* 213:527–545
9. Raudys S (2002) An adaptation model for simulation of aging process. *International Journal of Modern Physics C* 13:1075–1086
10. Alders P, Broer P (2005) Ageing, fertility, and growth. *Journal of Public Economics* 89:1075–1095
11. Thatcher AR (1999) The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society Series A (Statistics in Society)* 162:5–43
12. Bongaarts J, Feeney G (2002) How long do we live? *Population and Development Review* 28:13–29
13. Bongaarts J, Feeney G (2003) Estimating mean lifetime. *Proc Natl Acad Sci* 100:13127–13133
14. Lee ET (1992) *Statistical Methods for Survival Data Analysis*. Wiley, New York
15. Lee R, Carter LR (1992) Modeling and forecasting U.S. mortality. *J Am Statist Assoc* 87:659–671
16. Higgins T (2003) Mathematical models of mortality. Workshop on Mortality Modeling and Forecasting, Australian National University, 13–14 February
(<http://acpr.edu.au/Publications/Mortality%20talk%20-%20mathematical%20equations.pdf>)
17. Pletcher S (1996) Age-specific mortality costs of exposure to inbred *Drosophila melanogaster* in relation to longevity selection. *Experimental Gerontology* 31:605–616
18. Pletcher SD (1999) Model fitting and hypothesis testing for age-specific mortality data. *Journal of Evolutionary Biology* 12:430–439

19. Gompertz B (1825) On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* 115:513–585
20. Pham H (2002) A vtub-shaped hazard rate function with applications to system safety. *International Journal of Reliability Applications* 3:1–16
21. Weibull W (1951) A statistical distribution function of wide applicability. *Journal of Applied Mechanics* 18:293–297
22. Pham H, Lai CH (2007) On recent generalization of the Weibull distribution. *IEEE Transactions on Reliability* 56:454–458
23. Yu V, Pakin S, Hrisanov M (1984) Critical analysis of the applicability of the Gompertz–Makeham law in human populations. *Gerontology* 30:8–12
24. Bongaarts J (2005) Long-range trends in adult mortality: models and projection methods. *Demography* 42:23–49
25. Booth H, Maingdonald J, Smith L (2002) Applying Lee–Carter under conditions of variable mortality decline. *Population Studies* 56:325–336
26. Carey JR (1993) *Applied Demography for Biologists*. Oxford University Press
27. Finch CE (1990) *Longevity, Senescence and the Genome*. University of Chicago Press
28. Fisher RA (1921) On the mathematical foundation of theoretical statistics. *Philosophical Transactions of the Royal Society of London A* 222:309–368
29. Gavrilov LA, Gavrilova NS (1991) *The Biology of Life Span: a Quantitative Approach*. Harwood Academic Publishers, Switzerland
30. Hughes KA, Charlesworth B (1994) A generic analysis of senescence in *Drosophila*. *Nature* 367:64–66
31. Kannisto V (1987) On the survival of centenarians and the span of life. *Population Studies* 42:389–406
32. Kannisto V, Lauritsen J, Thatcher AR, Vaupel JW (1994) Reductions in mortality at advanced ages: several decades of evidence from 27 countries. *Population and Development Review* 20:793–810
33. Keyfitz N (1977) *Applied Mathematical Demography*. Wiley, New York
34. Keyfitz N (1981) Choice of function for mortality analysis: effective forecasting depends on a minimum parameter representation. *Theoretical Population Biology* 21:329–352
35. Riggs JE (1990) Longitudinal Gompertzian analysis of adult mortality in the US, 1900–1986. *Mechanisms of Ageing and Development* 54:235–247
36. Riggs JE (2001) The dynamic of aging and mortality in the United States. *Mechanisms of Ageing and Development* 65:217–228
37. Thatcher AR, Kannisto V, Vaupel JW (1998) *The Force of Mortality at Ages 80 to 120*. Odense University Press, Odense
38. Brooks A, Lithgow GJ, Johnson TE (1994) Mortality rates in a genetically heterogeneous population of *Caenorhabditis elegans*. *Science* 263:668–671
39. McNown R, Rogers A (1989) Forecasting mortality: a parameterized time series approach. *Demography* 26:654–660
40. Olshansky SJ (1988) On forecasting mortality. *The Milbank Quarterly* 66:482–530
41. Pollard JH (1987) Projection of age-specific mortality rates. *Population Bulletin of the United Nations*, nos 21–22. United Nations, New York, pp. 55–69
42. Mueller LD, Nusbaum TJH, Rose MR (1995) The Gompertz equation as a predictive tool in demography. *Experimental Gerontology* 30:553–569
43. Wilson DL (1993) A comparison of methods for estimating mortality parameters from survival data. *Mechanisms of Ageing and Development* 66:269–281
44. Wilson DL (1994) The analysis of survival (mortality) data: fitting Gompertz, Weibull and logistic functions. *Mechanisms of Ageing and Development* 74:15–33
45. Pakin YV, Hrisanov SM (1984) Critical analysis of the applicability of the Gompertz–Makeham law in human populations. *Gerontology* 30:8–12
46. Lee R, Miller T (2001) Evaluating the performance of the Lee–Carter method for forecasting mortality. *Demography* 38:537–549

47. Manton KG, Stallard E, Vaupel JW (1981) Methods for comparing the mortality experience of heterogeneous populations. *Demography* 18:389–410
48. Manton KG, Stallard E, Vaupel JW (1986) Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association* 81:635–644
49. Nusbaum TJ, Mueller LD, Rose MR (1996) Evolutionary patterns among measures of aging. *Experimental Gerontology* 31:507–516
50. Taljapourkar S, Li N, Boe C (2000) A universal pattern of mortality decline in the G7 countries. *Nature* 405:789–792
51. Curtsinger JW, Fukui HH, Townsend DR, Vaupel JW (1992) Demography of genotypes: failure of the limited lifespan paradigm in *Drosophila melanogaster*. *Science* 258:461–463
52. Beard RE (1971) Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In Brass W (ed) *Biological Aspects of Demography*. Barnes & Noble Inc., New York
53. Himes CL, Preston SH, Condran GA (1994) A relational model of mortality at older ages in low mortality countries. *Population Studies* 48:269–291
54. Horiuchi C, Coale AJ (1994) Age patterns of mortality for older women: an analysis using the age-specific rate of mortality change with age. *Mathematical Population Studies* 2:245–267
55. Pletcher S, Curtsinger JW (1998) Mortality plateaus and the evolution of senescence: Why are mortality rates so low? *Evolution* 52:454–464
56. Shyamalkumar ND (2002) Analysis of mortality data using smoothing spline Poisson regression. Department of Statistics, University of Iowa
57. Tabeau EA, Heathcote C (2001) *Forecasting Mortality in Developed Countries*. Kluwer Academic Publishers, Dordrecht
58. Vaupel JW (1986) How change in age-specific mortality affects life expectancy. *Population Studies* 40:147–157
59. Yashin AI, Vaupel JW, Iachine IA (1994) A duality in aging: the equivalence of mortality models based on radically different concepts. *Mechanisms of Ageing and Development* 74:1–14
60. Alho JM (1990) Stochastic methods in population forecasting. *International Journal of Forecasting* 6:521–530
61. Carey JR, Liedo P, Orozco D, Vaupel JW (1992) Slowing of mortality rates at older ages in large medfly cohorts. *Science* 258:457–461

Further Reading

Lee R (2000) The Lee–Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal* 4:80–93



<http://www.springer.com/978-1-84800-112-1>

Recent Advances in Reliability and Quality in Design

Pham, H. (Ed.)

2008, XXIV, 524 p., Hardcover

ISBN: 978-1-84800-112-1