U.S. Mortality at Advanced Ages: Cohort Analysis

Leonid A. Gavrilov
Natalia S. Gavrilova

Center on Aging
NORC and The University of Chicago
Chicago, Illinois, USA

The growing number of persons living beyond age 80 underscores the need for accurate measurement of mortality at advanced ages.

Recent projections of the U.S. Census Bureau significantly overestimated the actual number of centenarians

---

Views about the number of centenarians in the United States 2009

Centenarians are the fastest-growing age segment: Number of 100-year-olds to hit 6 million by 2050

BY THE ASSOCIATED PRESS

TUESDAY, JULY 21, 2009, 10:27 AM

New estimates based on the 2010 census are two times lower than the U.S. Bureau of Census forecast

Far fewer centenarians than expected in Census

1

Print Email Share Tweet Like Pin it

Posted Sept. 24, 2011, at 6:19 a.m.
Last modified Sept. 24, 2011, at 7:06 a.m.

NEW YORK — Reports of Americans living beyond the ripe old age of 100, it appears, were greatly exaggerated.

The Census Bureau predicted six years ago that the country would be home to 114,000 centenarians by 2010. The actual number was 53,364, the census reported recently. That represented an increase of 5.8 percent since 2000, compared with a 9.7 percent gain in the nation's population as a whole.

The same story recently happened in the Great Britain

Financial Times

September 11, 2012 8:20 pm

Long-lived Britons increasing slower than forecast

By Norma Cohen, Economics Correspondent

The rate at which Britons are living into very old age is rising much more slowly than had been forecast only two years ago, a blow for those hoping for a very long life but good news for pension providers and the Treasury which spend hefty sums on the oldest old.

Mortality at advanced ages is the key variable for understanding population trends among the oldest-old


THE WALL STREET JOURNAL

Death Gets in the Way of Old-Age Gains

Survivors

The first comprehensive study of mortality at advanced ages was published in 1939

HUMAN BIOLOGY

a record of research

FEBRUARY, 1939

VOL. 11  No. 1

THE BIOSTATISTICS OF SENILITY

BY MAJOR GREENWOOD AND J. O. IRWIN

M. Greenwood, J. O. Irwin. BIOSTATISTICS OF SENILITY

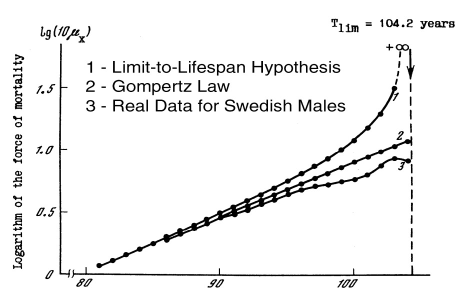
"the increase of mortality rate with age advances at a slackening rate, that nearly all, perhaps all, methods of graduation of the type of Gompertz's formula *over-state* senile mortality."

"... *possibility* that with advancing age the rate of mortality asymptotes to a finite value."

"... The limiting values of q_x are 0.439 for women and 0.544 for men. Some tests of the ultimate mortalities in non-human experience were not unfavorable."

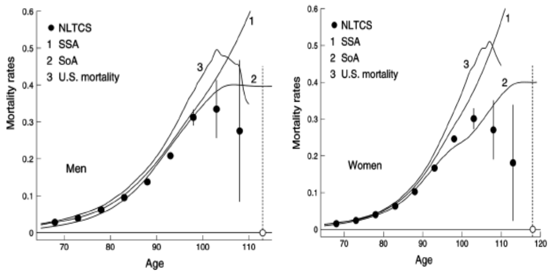
Earlier studies suggested that the exponential growth of mortality with age (Gompertz law) is followed by a period of deceleration, with slower rates of mortality increase.

Mortality at Advanced Ages – over 20 years ago



Source: **Gavrilov L.A., Gavrilova N.S. The Biology of Life Span: A Quantitative Approach, NY: Harwood Academic Publisher, 1991**

Mortality at Advanced Ages, Recent Study



Source: Manton et al. (2008). Human Mortality at Extreme Ages: Data from the NLTCS and Linked Medicare Records. *Math.Pop.Studies*

Existing Explanations of Mortality Deceleration

- **Population Heterogeneity (Beard, 1959; Sacher, 1966).** *"… sub-populations with the higher injury levels die out more rapidly, resulting in progressive selection for vigour in the surviving populations"* **(Sacher, 1966)**
- **Exhaustion of organism's redundancy (reserves) at extremely old ages so that every random hit results in death (Gavrilov, Gavrilova, 1991; 2001)**
- **Lower risks of death for older people due to less risky behavior (Greenwood, Irwin, 1939)**
- **Evolutionary explanations (Mueller, Rose, 1996; Charlesworth, 2001)**

Mortality force (hazard rate) is the best indicator to study mortality at advanced ages

$$\mu_x = -\frac{dN_x}{N_x dx} = -\frac{d\ln(N_x)}{dx} \approx -\frac{\Delta\ln(N_x)}{\Delta x}$$

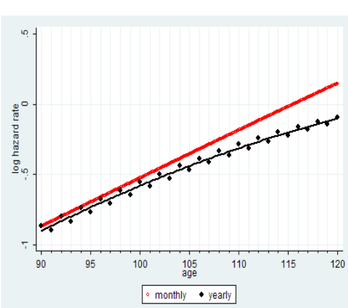
- Does not depend on the length of age interval
- Has no upper boundary and theoretically can grow unlimitedly
- Famous Gompertz law was proposed for fitting age-specific mortality force function (Gompertz, 1825)

Problems in Hazard Rate Estimation At Extremely Old Ages

1. Mortality deceleration in humans may be an artifact of mixing different birth cohorts with different mortality (heterogeneity effect)
2. Standard assumptions of hazard rate estimates may be invalid when risk of death is extremely high
3. Ages of very old people may be highly exaggerated

Monthly Estimates of Mortality are More Accurate

Simulation assuming Gompertz law for hazard rate



Stata package uses the Nelson-Aalen estimate of hazard rate:

$$\mu_x = H(x) - H(x-1) = \frac{d_x}{n_x}$$

H(x) is a cumulative hazard function, d_x is the number of deaths occurring at time x and n_x is the number at risk at time x before the occurrence of the deaths. This method is equivalent to calculation of probabilities of death:

$$q_x = \frac{d_x}{l_x}$$

Social Security Administration's Death Master File (SSA's DMF) Helps to Alleviate the First Two Problems

- Allows to study mortality in large, more homogeneous single-year or even single-month birth cohorts
- Allows to estimate mortality in one-month age intervals narrowing the interval of hazard rates estimation

What Is SSA's DMF ?

- As a result of a court case under the Freedom of Information Act, SSA is required to release its death information to the public. SSA's DMF contains the complete and official SSA database extract, as well as updates to the full file of persons reported to SSA as being deceased.
- SSA DMF is no longer a publicly available data resource (now is available from Ancestry.com for fee)
- We used DMF full file obtained from the National Technical Information Service (NTIS). Last deaths occurred in September 2011.

SSA's DMF Advantage

- Some birth cohorts covered by DMF could be studied by the method of extinct generations
- Considered superior in data quality compared to vital statistics records by some researchers

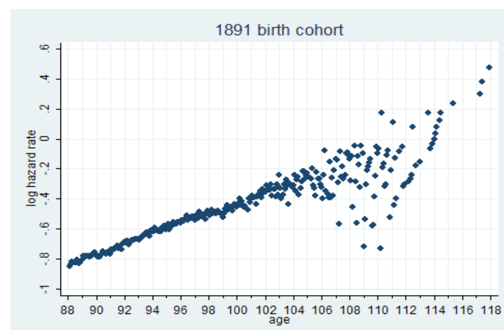
Social Security Administration's Death Master File (DMF) Was Used in This Study:

To estimate hazard rates for relatively homogeneous single-year extinct birth cohorts (1890-1899)

To obtain monthly rather than traditional annual estimates of hazard rates

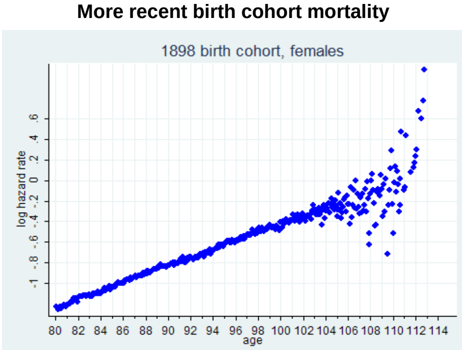
To identify the age interval and cohort with reasonably good data quality and compare mortality models

Hazard rate estimates at advanced ages based on DMF



Nelson-Aalen monthly estimates of hazard rates using Stata 11

More recent birth cohort mortality



Nelson-Aalen monthly estimates of hazard rates using Stata 11

Hypothesis

Mortality deceleration at advanced ages among DMF cohorts may be caused by poor data quality (age exaggeration) at very advanced ages

If this hypothesis is correct then mortality deceleration at advanced ages should be less expressed for data with better quality

Quality Control (1)

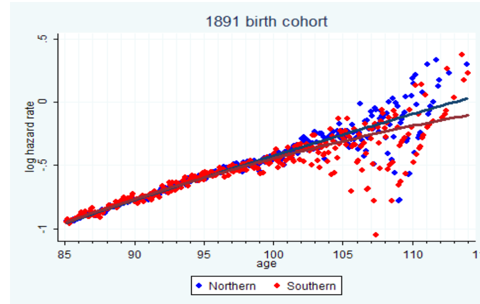
Study of mortality in the states with different quality of age reporting:

Records for persons applied to SSN in the Southern states were found to be of lower quality (Rosenwaike, Stone, 2003)

We compared mortality of persons applied to SSN in Southern states, Hawaii, Puerto Rico, CA and NY with mortality of persons applied in the Northern states (the remainder)

Mortality for data with presumably different quality:

Southern and Non-Southern states of SSN receipt



The degree of deceleration was evaluated using quadratic model

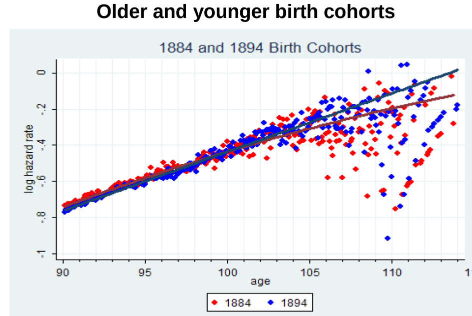
Quality Control (2)

Study of mortality for earlier and later single-year extinct birth cohorts:

Records for later born persons are supposed to be of better quality due to improvement of age reporting over time.

Mortality for data with presumably different quality:

Older and younger birth cohorts

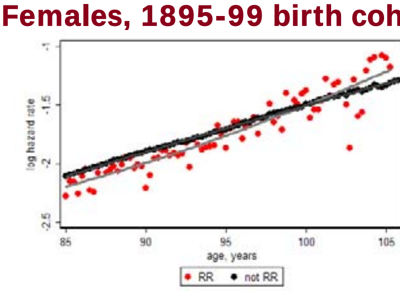


The degree of deceleration was evaluated using quadratic model

New Pilot Study based on DMF: Mortality of Railroad Retirees (SSN: 700-728)

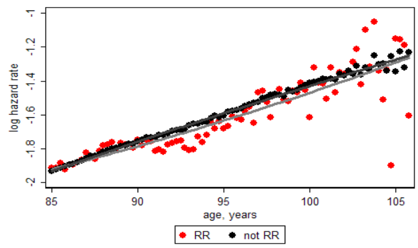
- In the past railroad workers could have better age reporting compared to their peers
- If mortality deceleration is caused by age misreporting, then better data quality for railroad workers may lead to less mortality deceleration among them

Mortality of Railroad Retirees and their non-Railroad Peers Females, 1895-99 birth cohort



Straight lines correspond to the quadratic fit of hazard rates in semi-log coordinates. For RR group, coefficient at quadratic term is positive and significant; for not-RR group this coefficient is not significant

Mortality of Railroad Retirees and their non-Railroad Peers Males, 1895-99 birth cohort

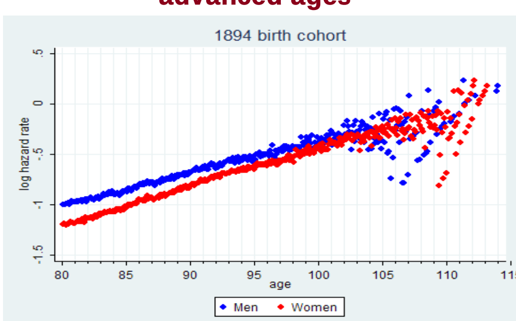


Straight lines correspond to the quadratic fit of hazard rates in semi-log coordinates
For RR group, coefficient at quadratic term is positive and significant; for not-RR group this coefficient is not significant

At what age interval data have reasonably good quality?

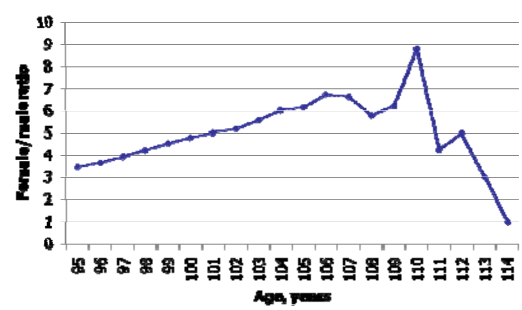
A study of age-specific mortality by gender

Women have lower mortality at advanced ages

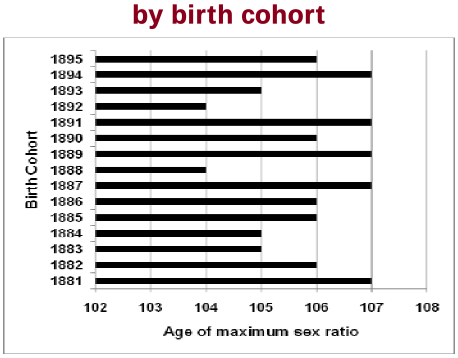


Hence number of females to number of males ratio should grow with age

Observed female to male ratio at advanced ages for combined 1887-1892 birth cohort



Age of maximum female to male ratio by birth cohort



Rapid deterioration of data quality with age in SSA DMF

Research Article
Typologies of Extreme Longevity Myths

Robert D. Young,[1] Bertrand Desjardins,[2] Kirsten McLaughlin,[1] Michel Poulain,[3] and Thomas T. Perls[1]

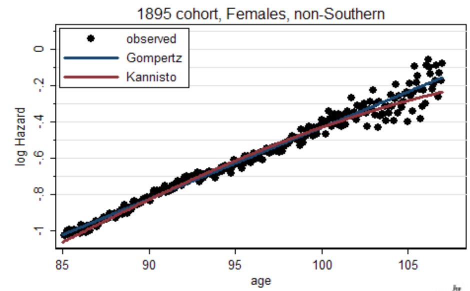
- "Invalid age claim rates increase with age from 65% at age 110-111 to 98% by age 115 to 100% for 120+ years."

(*Current Gerontology and Geriatrics Research*, Vol.2010, Article ID 423087)

Selection of competing mortality models using DMF data

- Data with reasonably good quality were used: non-Southern states and 85-106 years age interval
- Gompertz and logistic (Kannisto) models were compared
- Nonlinear regression model for parameter estimates (Stata 11)
- Model goodness-of-fit was estimated using AIC and BIC

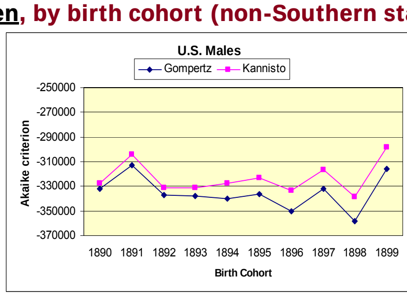
Fitting mortality with Kannisto and Gompertz models



Gompertz model $\mu_x = ae^{bx}$.

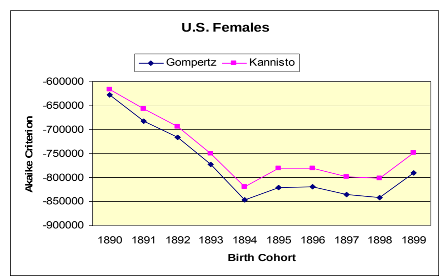
Kannisto model $\mu_x = \frac{ae^{bx}}{1+ae^{bx}}$

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **men**, by birth cohort (non-Southern states)



Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 85-106 years

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **women**, by birth cohort (non-Southern states)



Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 85-106 years

The second studied dataset: U.S. cohort death rates taken from the Human Mortality Database

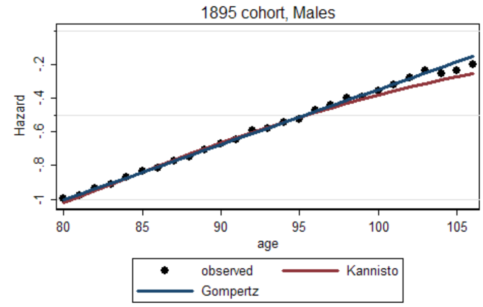
What is the Human Mortality Database?

- The Human Mortality Database (HMD) was created to provide detailed mortality and population data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity.
- URL: **http://www.mortality.org/**

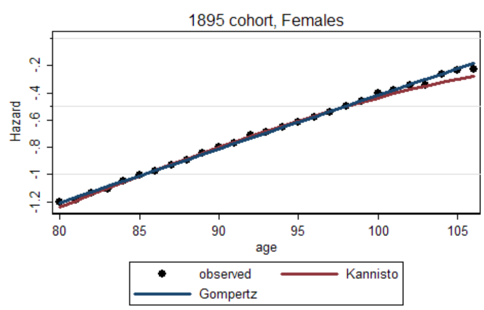
Selection of competing mortality models using HMD data

- Data with reasonably good quality were used: 80-106 years age interval
- Gompertz and logistic (Kannisto) models were compared
- Nonlinear weighted regression model for parameter estimates (Stata 11)
- Age-specific exposure values were used as weights (Muller at al., Biometrika, 1997)
- Model goodness-of-fit was estimated using AIC and BIC

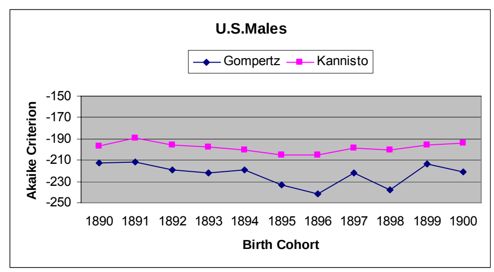
Fitting mortality with Kannisto and Gompertz models, HMD U.S. data



Fitting mortality with Kannisto and Gompertz models, HMD U.S. data

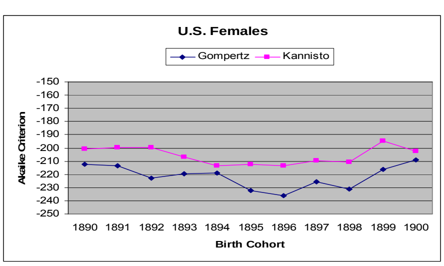


Akaike information criterion (AIC) to compare Kannisto and Gompertz models, <u>men</u>, by birth cohort (HMD U.S. data)



Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 80-106 years

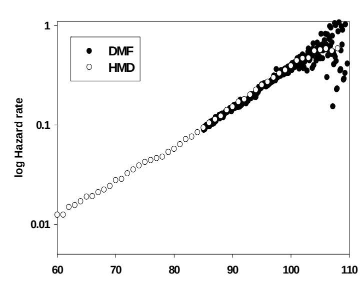
Akaike information criterion (AIC) to compare Kannisto and Gompertz models, <u>men</u>, by birth cohort (HMD U.S. data)



Conclusion: In all ten cases Gompertz model demonstrates better fit than logistic model for men in age interval 80-106 years

Compare DMF and HMD data

Females, 1898 birth cohort



Hypothesis about two-stage Gompertz model is not supported by real data

Which estimate of hazard rate is the most accurate?

Simulation study comparing several existing estimates:

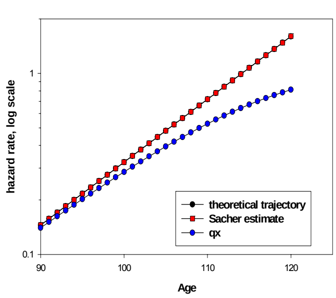
- **Nelson-Aalen estimate available in Stata**
- **Sacher estimate (Sacher, 1956)**
- **Gehan (pseudo-Sacher) estimate (Gehan, 1969)**
- **Actuarial estimate (Kimball, 1960)**

Simulation study to identify the most accurate mortality indicator

- Simulate yearly l_x numbers assuming Gompertz function for hazard rate in the entire age interval and initial cohort size equal to 10^{11} individuals
- Gompertz parameters are typical for the U.S. birth cohorts: slope coefficient (alpha) = 0.08 $year^{-1}$; R_0= 0.0001 $year^{-1}$
- Focus on ages beyond 90 years
- Accuracy of various hazard rate estimates (Sacher, Gehan, and actuarial estimates) and probability of death is compared at ages 100-110

Simulation study of Gompertz mortality

Compare Sacher hazard rate estimate and probability of death in a yearly age interval



Sacher estimates practically coincide with theoretical mortality trajectory

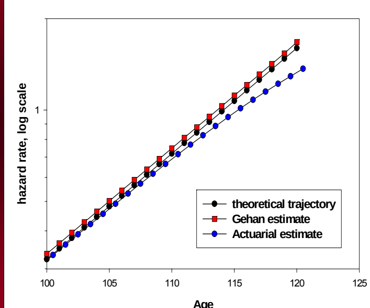
$$\mu_x = \frac{1}{2\Delta x} ln\frac{l_{x-\Delta x}}{l_{x+\Delta x}}$$

Probability of death values strongly undeestimate mortality after age 100

$$q_x = \frac{d_x}{l_x}$$

Simulation study of Gompertz mortality

Compare Gehan and actuarial hazard rate estimates



Gehan estimates slightly overestimate hazard rate because of its half-year shift to earlier ages

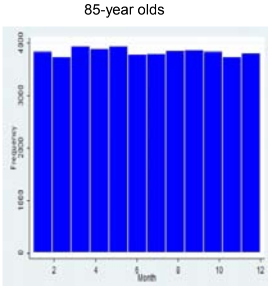
$$\mu_x = -\ln(1 - q_x)$$

Actuarial estimates undeestimate mortality after age 100

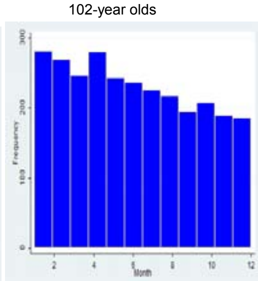
$$\mu_{x+\frac{\Delta x}{2}} = \frac{2}{\Delta x}\frac{l_x - l_{x+\Delta x}}{l_x + l_{x+\Delta x}}$$

Deaths at extreme ages are not distributed uniformly over one-year interval

85-year olds



102-year olds



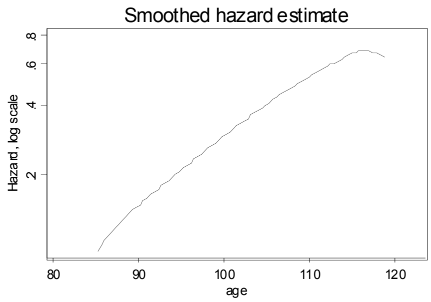
1894 birth cohort from the Social Security Death Index

Accuracy of hazard rate estimates

Relative difference between theoretical and observed values, %		
Estimate	**100 years**	**110 years**
Probability of death	11.6%, understate	26.7%, understate
Sacher estimate	0.1%, overstate	0.1%, overstate
Gehan estimate	4.1%, overstate	4.1%, overstate
Actuarial estimate	1.0%, understate	4.5%, understate

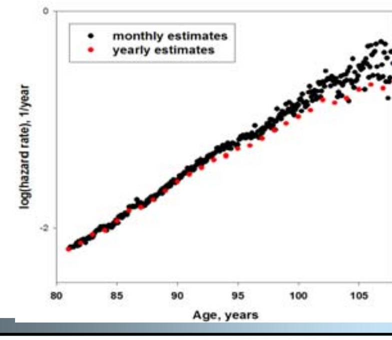
A word of caution: Data smoothing may lead to spurious mortality deceleration

What happens with simulated Gompertz law data after their Kernel smooth by Stata (default settings):



Mortality of 1894 birth cohort

Monthly and Yearly Estimates of Hazard Rates using Nelson-Aalen formula (Stata)



Sacher formula for hazard rate estimation (Sacher, 1956; 1966)

$$\mu_x = \frac{1}{\Delta x}\left(\ln l_{x-\frac{\Delta x}{2}} - \ln l_{x+\frac{\Delta x}{2}}\right) = \frac{1}{2\Delta x}\ln\frac{l_{x-\Delta x}}{l_{x+\Delta x}}$$

Hazard rate

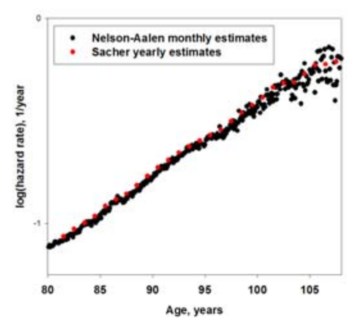
l_x - survivor function at age x; Δx – age interval

Simplified version suggested by Gehan (1969):

$$\mu_x = -\ln(1-q_x)$$

Mortality of 1894 birth cohort

Sacher formula for yearly estimates of hazard rates



Conclusions

- Deceleration of mortality in later life is more expressed for data with lower quality. Quality of age reporting in DMF becomes poor beyond the age of 107 years
- Below age 107 years and for data of reasonably good quality the Gompertz model fits mortality better than the logistic model (no mortality deceleration)
- Sacher estimate of hazard rate turns out to be the most accurate and most useful estimate to study mortality at advanced ages

Acknowledgments

This study was made possible thanks to:

generous support from the

- National Institute on Aging (R01 AG028620)
- Stimulating working environment at the Center on Aging, NORC/University of Chicago

For More Information and Updates Please Visit Our Scientific and Educational Website on Human Longevity:

- http://longevity-science.org

And Please Post Your Comments at our Scientific Discussion Blog:

- http://longevity-science.blogspot.com/