A discussion of the upper limit of human longevity based on study of data for oldest old survivors and deaths in Japan

Nobutane Hanayama
Shobi University, Kawagoe, Japan
e-mail: nob-hanayama@jcom.home.ne.jp

Abstracts

In the field of biology, theories of aging are roughly divided into two major groups. One is consisting of damage theories and the other is consisting of program theories. According to damage theories we age because our systems break down over time. Meanwhile, in the program theories, it is considered that we age because there is an inbuilt mechanism that tells us to die. If the damage theories are true, we can survive any longer by avoiding damaging our organism. If the program theories are true, on the other hand, we cannot survive longer than the upper limit of longevity with any effort. In this study, for discussing that either the damage theories (there exist a upper limit of human longevity) or the programing theories (there does not exist such a limit) is true, data for oldest old survivors and deaths given by age and birth-period in Japan is analyzed using the extreme value theory and the extinct cohort method. From the results of fitting the binomial regression model with probabilities calculated from the generalized Pareto distribution to the data, the upper limit of human longevity is estimated from 107 to 128 years for male or from 119 to 159 years based on the data for survivors, while that is estimated infinite for some cohorts or form 108 to 262 years for male or form 117 to 165 years for female based on the data for deaths.

Keywords: Age-by-period data, extreme value theory, generalized Pareto distribution

1. Introduction

In the field of biology theories of aging are roughly divided into two major groups. One is consisting of damage theories and the other is consisting of program theories. According to damage theories we age because our systems break down over time. Meanwhile, in the program theories, it is considered that we age because there is an inbuilt mechanism that tells us to die. If the damage theories are true, we can survive any longer by avoiding damaging our organism. If the program theories are true, on the other hand, we cannot survive longer than the upper limit of longevity with any effort.

Considering the above facts our aim in this study is to discuss whether the damage theories are true, which means that there exist a upper limit of human longevity, or the programing theories are true, which means that there does not exist such a limit, by using the extreme value theory in an analysis of data for Japanese oldest old survivors and deaths given by age and period (age-by-period data) which is introduced in the following Section 2 where the method of extinct cohort (Wilmoth et al., 2007) is introduced for figuring out the numbers of survivors at every age for every cohort from the data for deaths. In Section 3 our procedure for analyzing those data is introduced referring the extreme value theory and the results of the estimation of the upper limit of human longevity based on both survivors and deaths are shown in the following Section 4. In the final Section 5 the results of estimation of the upper limit of human longevity based on the data for survivors and deaths are discussed.

2. Data

In order to accomplish our aims, we need data for oldest old survivors, which are for survivors aged 100 years and older in general. In the report of national census, though the numbers of survivors are indicated in tables, data are aggregated by 5 year or totaled for 100 years and older in those tables. Thus the data obtained from the reports of national census is not available. So, in the followings two sets of data for oldest old survivors obtained from other sources than the national census are introduced.

2.1. National list of aged

The ministry of health, labors and welfare in Japan is (was) publishing a book which contains list of names, addresses and ages of people who are 100 years and older in every year. And, on the top page of those books, a table indicating the numbers of survivors is given by age. So, combining those tables for years, we can get a (age, period)-tabulated data for the numbers of survivors aged 100 years and older.

Unfortunately the National list of Aged has halted its publication since 2003 for protecting private information. In addition, a big scandal had been revealed in 2010: more than 230,000 elderly people in Japan who are listed as being aged 100 or over are unaccounted for, officials said following a nationwide inquiry on 10 September 2010. Because of that, the significance of those data has been vanished

2.2. Vital Statistics in Japan

The numbers of deaths by year when they were born, that is, by cohort are given in Vital Statistics in Japan. So, the population size for a cohort at a certain age is estimated by summing all future deaths for the cohort with the method of extinct cohort, which is one whose members are assumed to have all died by the end of the observation period.

The method of extinct cohort is explained as follows. Let $D_{ij}^{(U)}$ indicate the number of deaths who are at the age of i in the year j and were born in the year $j-i-1$ and $D_{ij}^{(L)}$ indicate the number of deaths who are at the age of i in the year j and were born in the year $j-i$. Then, assuming that there is no migration or error, the number of survivors who have reached the age of i, say S_{ij}, is

$$S_{ij} = D_{i,j}^{(L)} + \sum_{k=1}^{\infty}\left(D_{i+k,j+k}^{(U)} + D_{i+k+1,j+k}^{(L)}\right). \qquad (1)$$

The method of extinct cohort is described also in Methods Protocol for the Human Mortality Database by Wilmoth, et al. (2007).

3. Procedure

Let $X_1, X_2, \cdots, X_n$ be a sequence of independent lifetime of individuals with common distribution function F and denote an arbitrary term in the X_i sequence by X. Then, according to the extreme value theory, if it can be considered that $u > 100$ is a large enough constant, the distribution function of $X-u$ conditional on $X > u$ is approximately the generalized Pareto distribution, which is written as

$$F(y;\gamma,a) = \begin{cases} 1-(1+\gamma y/a)^{-1/\gamma} & \gamma \neq 0 \\ 1-\exp(-y/a) & \gamma = 0\,. \end{cases} \qquad (2)$$

An important property of this distribution is that the upper limit of distribution varies depending on parameter values. That is, if $\gamma < 0$ then the upper limit of the distribution is finite, that is $0 < Y < -a/\gamma \equiv \omega$, and if $\gamma \geq 0$, then the upper limit of the distribution is infinite, that is, $0 < Y < \infty$. Considering those properties, a procedure for discussing the limit of human longevity is this study is laid out as follows:

1. to test the hypothesis $\gamma = 0$,
2. to estimate the upper limit of life for each cohort,

The generalized Pareto distribution is one for continuous random variable, meanwhile our data is aggregated by year and given as discrete data. So, we need to associate continuous survival time with discrete data. For that purpose, the Lexis diagram (Keiding, 1990) is a useful tool (see figure 1). In the figure the horizontal axis indicates the time and the vertical axis indicates the age and oblique upward lines indicate individuals who are alive. And, those vertical lines indicate some fixed days or time points in each year. Then the numbers of survivors who are $100+i$ years old ($i=1,2,\cdots,I$) in the year $1976+j$ ($j=1,2,\cdots,26$), say M_{ij}, equals the counts of cross points of oblique lines and vertical lines, where I is supposed to be sufficiently large so that $M_{Ij}=0$ for all j, and the numbers of deaths in a year are obtained as difference of M_{ij}, that is, $m_{ij} = M_{i+1,j+1} - M_{ij}$ (see Figure 2).

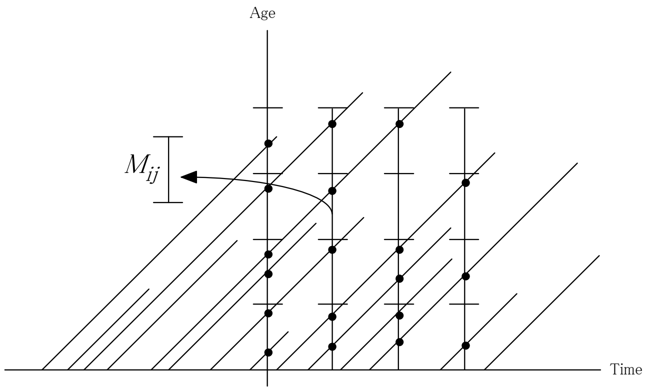

Figure 1. Lexis diagram and the number of survivors at every time point (year)

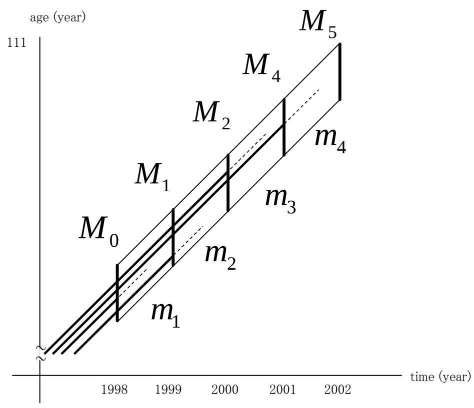

Figure 2. Numbers of deaths in a year are obtained as difference of M_{ij}

Now assume that individuals die in the year $1976+j$ at the age of $100+i$ with probabilities given by $p_j = F(i+1,\gamma_j,a_j) - F(i,\gamma_j,a_j)$. Then the contribution of m_{ij} to the likelihood is $\left(M_{ij}/m_{ij}\right) \; q_{ij}^{m_{ij}}(1-q_{ij})^{M_{ij}-m_{ij}}$, where $q_{ij} = p_j/(1-F(i,\gamma_j,a_j))$, hence the likelihood is proportional to

$$\prod_{i=1,j=1}^{I,J} \binom{M_{ij}}{m_{ij}} q_{ij}^{m_{ij}} (1-q_{ij})^{M_{ij}-m_{ij}} . \qquad (3)$$

Thus models assumed on p_{ij} thus fall into the class of binomial regression models, and the parameters in the generalized Prato distribution are estimated by maximizing the likelihood function (3).

4. Results

Table 1 indicate the ML estimates of the parameter γ and the upper limit of life, that is ω, based on the data for male and female survivors respectively. As shown in Table 1, the estimates of γ are negative for all cohorts of male and female, though some p-values are sufficiently small while the others are not so. Besides, the upper limit of life, that is ω, is estimated between 107 and 128 though the hypothesis $\gamma = 0$ is not rejected for the years 1967and 1972 as indicated in Table 1 for male, while those are estimated between 116 and 161 for females while all p-values are sufficiently small as indicated in Table 2.

Table 1. Estimates based on the data for survivors for male (left) and female (right)

Year	Gamma	P-value (one side)	Omega	P-value (one side)
1967	0.08	0.84		
1972	0.11	0.92		
1977	-0.30	0.00	107.00	0.00
1982	-0.14	0.00	115.50	0.00
1987	-0.09	0.00	121.56	0.00
1992	-0.11	0.00	119.82	0.00
1997	-0.08	0.01	128.00	0.00

Year	Gamma	P-value (one side)	Omega	P-value (one side)
1967	-0.03	0.23	159.33	0.00
1972	-0.03	0.24	161.67	0.00
1977	-0.13	0.00	116.92	0.00
1982	-0.12	0.00	119.42	0.00
1987	-0.1	0.00	123.40	0.00
1992	-0.11	0.00	123.09	0.00
1997	-0.16	0.00	117.81	0.00

Table 2 indicates the ML estimates of the parameter γ and ω obtained by applying the method of extinct cohort to the data for male and female deaths. As shown in Table 2, the estimates of γ are negative for all male cohorts with sufficiently small p-values except the cohort 1968 for male and the cohorts 1944, 1954 and 1959 for female. Besides, the upper limit of life, that is ω, is estimated between 115 and 262, while those are estimated between 117 and 143 for females with sufficiently small p-values.

Table 2. Estimates for male (left) and female (right) obtained by applying the method of extinct cohort to the data for deaths

Year	Gamma	P-value (one side)	Omega	P-value (one side)
1948	-0.16	0.08	117.06	0.00
1953	-0.13	0.13	115.77	0.00
1958	-0.22	0.01	108.41	0.00
1963	0.12	0.88		
1968	-0.08	0.09	127.38	0.00
1973	-0.01	0.43	262.00	0.00
1978	-0.07	0.04	129.43	0.00
1983	-0.09	0.00	123.44	0.00
1988	-0.13	0.00	116.62	0.00
1993	-0.10	0.00	122.40	0.00

Year	Gamma	P-value (one side)	Omega	P-value (one side)
1944	0.18	0.85		
1949	-0.07	0.16	128.71	0.00
1954	0.07	0.88		
1959	0.03	0.70		
1964	-0.09	0.01	121.78	0.00
1969	-0.12	0.00	117.83	0.00
1974	-0.03	0.11	165.00	0.00
1979	-0.05	0.00	143.40	0.00
1984	-0.12	0.00	120.00	0.00
1989	-0.12	0.00	120.83	0.00

4. Discussion

In the field of demography, human lifetime distributions are conventionally assumed to be Gompertz curves, with the semi-infinite range, being fitted to the whole human lifetime, or to the adult lifetime (see Kanisto, 1999). Hence, fitting Generalized Pareto model with a finite upper limit to human lifetime demonstrates an alternative to the traditional convention.

Kaufmann（2001） has analyzed mortality data from West Germany and suggested an improbability that there exist an limit of human longevity. On the other hand Arssen and de Haan (1994) studied the same problem using datasets of both birthdays and death days. Their data consists of the total life span (in days), sex, and year of birth of all people born in the years 1877-1881, still alive on January 1, 1971 and who died as residents of the Netherlands. The size of table is 4131 men and 6260 women. Their results are consistent with that of the present paper. They showed that there is a finite age limit in 113–124 years. Our results of fitting the generalized Pareto model to the data for survivors and deaths have not supported either of their results.

In addition our estimates for the limit of human longevity have fluctuated depends on cohorts. From those results it has been considered that more detailed study is needed.

References

Arssen, K., de Haan, L. (1994). On the maximal life span of humans. Mathematical Population Studies, Vol. 4, 259-281.

Kannisto, V. (1999). Trends in the mortality of the oldest-old, Statistics, Registries, and Science: Experiences from Finland (ed. Alho, J.), Statistics Finland, 177-194.

Keiding, N. (1990). Statistical inference in the Lexis diagram. Philosophical Transactions of the Royal Society of London Series A, Vol. 332, 487-509.

Wilmoth, J.R., Andreev, K., Jdanov, D. and Glei, D.A. (2007). Methods Protocol for the Human Mortality Database. Vachonhttp://www.mortality.org/Public/Docs/ MethodsProtocol.pdf.

Acknowledgment. This work was supported by Grant-in-Aid for Scientific Research 24500346.