

ESTIMATING THE HEALTH STATUS OF A POPULATION: HEALTH STATE CURVES AND SURVIVAL CURVES

Christos H Skiadas

Abstract

In this paper we explore the very important case of finding a health measure in the lines of the survival curve but independent of the standard deviation parameter σ . This is done by estimating the health state curve and calculating the total area under the curve. An interesting comparison with the survival curves is done. The health status measure (HSM) resulting is of the form of the life expectancy (LE) as it is expressed in terms of years of age. The HSM is independent of the standard deviation σ and coincides with the LE for $\sigma=0$, that is when a perfect rectangularization of the survival curve appear. The provided HSM form gives rise to an interesting classification of various countries. The results are close to those obtained by the LE estimator with an interesting reorganization of country ranking.

Key words: Health status, health state, survival curve, standard deviation, life expectancy.

JEL Code: I15, I18, J1

1 Introduction

The form of the survival curve resulting from a classical life table is a measure of the health state or health status of a population. Much attention has given to the exploration of the part of the influence of disability on the form of the survival curve thus providing methods for estimating the loss of healthy life years in a population. However, the form of the survival curve depends on the dispersion of deaths around a mean value, measured by the standard deviation σ thus making the survival curve not very appropriate when doing comparisons between different populations.

An approach based on a simple Inverse Gaussian model and applied to Carey Medfly data is due to Weitz and Fraser (2001). The Inverse Gaussian form cannot apply to the human population data. The needed general form is given in Skiadas and Skiadas (2010) whereas the theoretical and technical issues are presented in Janssen and Skiadas (1995).

The main points in estimating the HSM are: to introduce a “system of measure” a “metric” that is independent of the standard deviation σ , to effectively estimate the health state, to

reproduce the population behavior, to compare the health state in various countries and populations for the same or various time periods, to be able to calculate to cost of the health improvement of the population, to estimate the insurance cost and the insurance policies.

1.1 The Related Theory

Following the first exit time theory we assume that the health state of an individual is expressed by a stochastic function denoted by S_x and the associated stochastic paths over time t or age x are estimated after integrating the stochastic differential equation

$$dS_x = h_x dx + \sigma dW_x \quad (1)$$

With drift h_x and finding the formula for the stochastic paths S_x

$$S_x = \int_0^x h_s ds + \int_0^x dW_s = H_x + \sigma W_x \quad (2)$$

Where W_x is the Wiener process and the Health State H_x is provided as the integral of the instantaneous change h_x

$$h_x = \frac{dH_x}{dx} \quad (3)$$

The death occurs when $S_x = 0$ and from (2) follows that

$$H_x + \sigma W_x = 0 \quad (4)$$

The simpler form for the Health State H_x should be a decreasing process of the form (see related bibliography in Janssen and Skiadas 1995 and Skiadas 2010, 2013, 2014, 2015):

$$H_x = l - (bx)^c \quad (5)$$

Where l , b and c are parameters. The form of (4) becomes

$$l - (bx)^c + \sigma W_x = 0 \quad (6)$$

This is a very important relation providing different ways of simulation of the stochastic process.

This is demonstrated by observing the new form of (2) that is

$$S_x = l - (bx)^c + \sigma W_x \quad (7a)$$

This form is important for constructing stochastic paths for the health state S_x .

The next form arises from (7a) by a simple transformation

$$S_x + (bx)^c = l + \sigma W_x \quad (7b)$$

A third form given by

$$S_x - l + (bx)^c = \sigma W_x \quad (7c)$$

The three forms (7a), (7b) and (7c) are mathematically the same. However, they provide three distinct simulation opportunities very important to explain the development of the health status and the development of the death probability density function.

According to the theory developed (Janssen and Skiadas, 1995) first it was solved the associated Fokker-Planck equation for the appropriate boundary conditions in order to find the transition probability density function and then we could find the formula for the first exit time probability density function $g(x)$ for the health state stochastic process crossing or hitting for the first time a barrier set to zero.

$$g(x) = \frac{|H_x - tH_x'|}{\sigma \sqrt{2\pi x^3}} e^{-\frac{(H_x)^2}{2\sigma^2 x}} \quad (8)$$

2 The Weitz and Fraser Paper Revisited

Fifteen years after the Weitz and Fraser (2001) paper only few applications appear though the high importance of this paper for the establishment of a system of measuring the health state or health status of a population. Perhaps the few applications are due to the paper title “Explaining mortality rate plateaus” instead of the use of the first exit time or hitting time theory in relation to the health status.

Weitz and Fraser introduced the simpler form for the health state, presented here by equation (5), but with the parameter $c = 1$. This is a form not convenient for the human population health status but it could be used to model the health status of a large group of Medflies systematically studied by Carey (1992).

$$H_x = l - (bx) \quad (9)$$

This form for the health status provides the following probability density function for the first exit time from a barrier set at zero level, that is:

$$g(x) = \frac{l}{\sigma \sqrt{2\pi x^3}} e^{-\frac{(l-bx)^2}{2\sigma^2 x}} \quad (10)$$

This is the classical form of the so-called Inverse Gaussian known at least from 1915.

When trying to estimate the parameters of (10) from the mortality data $d(x)$ or $g(x)$ by a non-linear regression analysis it is obvious that we cannot estimate simultaneously σ with the two parameters b and l . Instead by using the transformation $b^*=b/\sigma$ and $l^*=l/\sigma$ we arrive in the following (11) form providing the new parameters b^* and l^* as functions of σ . Instead Weitz and Fraser selected to set $l=1$ so that they estimated σ and b . Both methods can be useful for the estimates that follow. Note that setting a health state level equal to unity was proposed by Torrance (1976).

$$g(x) = \frac{(l/\sigma)}{\sqrt{2\pi x^3}} e^{-\frac{((l/\sigma)-(b/\sigma)x)^2}{2x}} = \frac{l^*}{\sqrt{2\pi x^3}} e^{-\frac{(l^*-b^*x)^2}{2x}} \quad (11)$$

2.1 Deterministic and Stochastic Case

An interesting question is related to the deterministic and stochastic case of the problem by means of estimating the mean value of the process. This is illustrated in Figure 1A where the mean value of the deterministic process is the line AB and the mean value of the stochastic process is precisely the same line AB. That is different is that the deterministic line is the result of the mean value of N members of the population identically distributed with $\sigma=0$ whereas, in the stochastic case, the mean value provided in line AB is the average of the stochastic paths of N members of the population. Although no-stochastic path is identical to another the mean value is identical to the deterministic case. That it is different is that $\sigma>0$. From (2) and (9) the stochastic paths are of the form:

$$S_x = H_x + \sigma W_x = l - bx + \sigma W_x \quad (12)$$

The deterministic case ($\sigma=0$) is given by:

$$S_x = H_x = l - bx \quad (13)$$

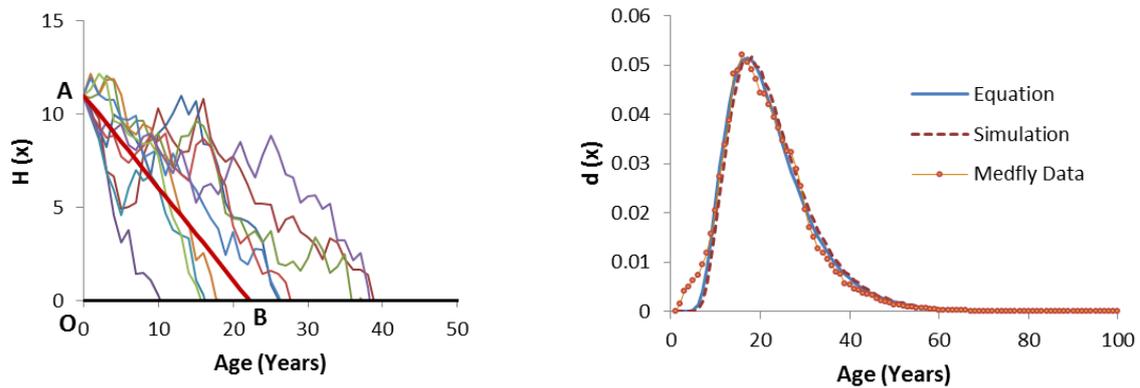
Following the deterministic case death occurs at $H_x=0$ at age $x=l/b$ (point B of Figure 1A). For the stochastic case the estimation procedure based on (11) will estimate $l^*=l/\sigma$ and $b^*=b/\sigma$. For this case death occurs at $H_x=0$ at age $x= l^*/ b^*=(l/\sigma)/(b/\sigma)=l/b$ that is precisely the same as in the deterministic case as it is independent of σ .

Note that the stochastic case provides the death probability density form of Figure 1B whereas the deterministic case will distribute all deaths exactly at point B of Figure 1A thus producing

a sharp distribution form at Figure 1B. The higher the standard deviation σ results in larger dispersion in the death distribution.

Instead the starting point of the process (the health state l at birth) is at level l for the deterministic process (point A of Figure 1A) whereas it is at level $l^*=l/\sigma$ for the stochastic process by means that the starting point for the stochastic process resulting from the estimation from (11) can be in a large variety of places in the Y axis depending on the selection of σ by means that the location of the point A is not definite and the same holds for the estimation of the level of the health state at birth. Instead it is perfectly located the zero health state at point B (Figure 1A).

Fig. 1. A. Stochastic paths and B. Stochastic simulation for Carey Medfly



For the deterministic case we can estimate the total health state estimated as the area of the triangle (OAB) that is $H_{total}=(OA)*(OB)/2=l*(l/b)/2=(l^2/b)/2$. The mean value is $H_{mean}=l/2$

For the stochastic case the total health state is $H_{total}=(OA)*(OB)/2=l^*(l/b)/2=(l/\sigma)(l/b)/2$. The mean value is $H_{mean}=l^*/2=(l/\sigma)/2$. This value is equal to the deterministic case only when $\sigma=1$.

Comparing different cases it is advisable to set $l^*=1$. The resulting mean health state level for the stochastic process is: $H_{mean}=l^*/2=(l/\sigma)/2=1/2$. The total health state is estimated as: $H_{total}=(OA)*(OB)/2=l^*(l^*/b^*)/2=(l^*/b^*)/2$.

The estimates of Weitz and Fraser for Carey Medfly data by the maximum likelihood estimation method give: $\sigma^*=0.0975$ corresponding to $l^*=1/\sigma^*=10.256$ and $b^*=0.448$.

Our estimates by a Levenberg–Marquardt nonlinear regression analysis algorithm provide: $l^*=10.532$ corresponding to $\sigma^*=1/l^*=0.0949$ and $b^*=0.480$.

The total health state is $H_{total}=(1/b^*)/2=11.446$ for Weitz and Frazer and 10.967 for our estimates.

The advanced method applying (5) for the same data sets provides the following parameter estimates:

$c^*=1.321$, $b^*=0.2370$ and $l^*=8.572$. The total health state is $H_{total}=12.211$.

2.2 The Main Points in Estimating the Health State of a Population

1. To introduce a “system of measure” a “metric” that is independent of the related situation
2. To effectively estimate the health state
3. To reproduce the population behavior by means of finding the death probability density, the life expectancy and other measures
4. To compare the health state in various countries and populations for the same or various time periods
5. To find a measure of the health state for all the age period by estimating a measure of the total health state as a summation or integration of all the period of the life span
6. To be able to calculate to cost of the health improvement of the population
7. To estimate the insurance cost and the insurance policies

3 The More General Case

The more general case for the health state (5) is of the form

$$H_x = l - (bx)^c \quad (14)$$

As in the previous case and using (8), this form for the health status provides the following probability density function for the first exit time from a barrier set at zero level, that is:

$$g(x) = \frac{|l + (c-1)(bx)^c|}{\sigma \sqrt{2\pi x^3}} e^{-\frac{(l-(bx)^c)^2}{2\sigma^2 x}} \quad (15a)$$

By using the transformation

$$b^*=b/\sigma^{1/c} \text{ and } l^*=l/\sigma \quad (15b)$$

we arrive in the following (16) form providing the new parameters b^* and l^* as functions of σ .

$$g(x) = \frac{|l/\sigma + (c-1)[(b/\sigma^{1/c} x)^c]|}{\sqrt{2\pi x^3}} e^{-\frac{(l/\sigma - (b/\sigma^{1/c} x)^c)^2}{2x}} = \frac{|l^* + (c-1)(b^* x)^c|}{\sqrt{2\pi x^3}} e^{-\frac{(l^* - (b^* x)^c)^2}{2x}} \quad (16)$$

The following alternative of (14) accepts the simpler transformation $b^*=b/\sigma$ and $l^*=l/\sigma$

$$H_x = l - bx^c \quad (17)$$

However, during the applications, the parameter b is very small causing problems in the fitting process.

Both (14) and (17) provide the same measures for the total health state. This is achieved by calculating the integral from age zero ($x=0$) to age $x=T$ of the zero health state.

$$H_{total} = \int_0^T l^* - (b^* x)^c dx = \left[xl^* - \frac{(b^*)^c x^{(c+1)}}{c+1} \right]_0^T = Tl^* \frac{c}{c+1} \quad (18)$$

Note that from (14) $H=0$ for $T=l^{*(1/c)}/b^*$.

The very important achievement is that by setting $l^*=1$ in (18) the total health state is given by:

$$H_{total} = T \frac{c}{c+1} = \frac{l^{*(1/c)}}{b^*} \cdot \frac{c}{c+1} \quad (19)$$

Note that the total health state is the more convenient estimator for the comparisons between different populations and countries. It is also in accordance with the main part of the life expectancy at birth ranking though a refined ranking emerges for some countries.

Accordingly the mean health state H_{mean} is given by:

$$H_{mean} = \frac{c}{c+1} \quad (20)$$

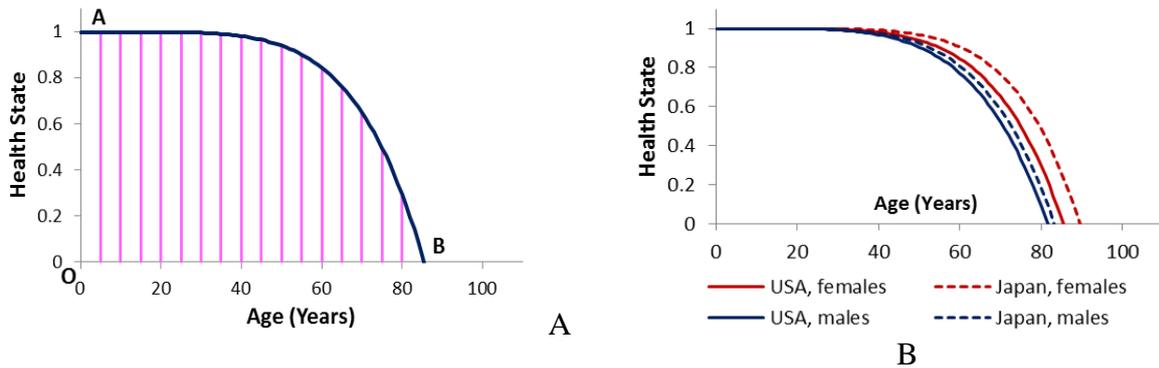
We can easily find that for the simple case ($c=1$) the mean health state is $H_{mean}=1/2$ and the total health state is $H_{total}=T/2=(l^*/b^*)/2$.

Note that the total health state is the more convenient estimator for the comparisons between different populations and countries. The main idea is coming from establishing a health state level equal to unity (see Figure 2A and 2B) at birth by setting $l^*=1$. This is an independent measure for every population and country. The next step is to find the total health state by evaluating the integral (18) providing the simple measure in (19). By using the transformation (15b) the right hand side of (19) is given by

$$H_{total} = T \frac{c}{c+1} = \frac{l^{*(1/c)}}{b^*} \cdot \frac{c}{c+1} = \frac{l^{(1/c)} / \sigma^{(1/c)}}{b / \sigma^{(1/c)}} \cdot \frac{c}{c+1} = \frac{l^{(1/c)}}{b} \cdot \frac{c}{c+1} \quad (21)$$

The very important achievement from (21) is that the total health state is independent of σ . When the exponent c is large the total health state is of the order of $l^{(1/c)}/b$. This is a simple method expressing the health state in terms of age as it is done for the life expectancy.

Fig. 2. A. Health State versus age and total health state area B. Comparing Health State



4 Connection of the health state curve to the survivorship curve

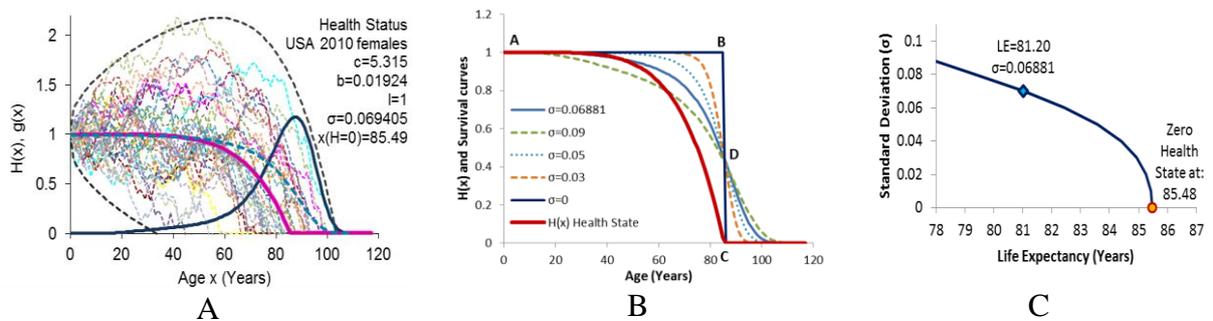
As it was presented earlier the health state curve is independent of σ . Instead the survivorship curve $l(x)$ or survival curve is connected to σ as can be verified by stochastic simulations. Even more the survivorship curve is connected with the death probability which includes σ as it expressed in (22).

$$l(x) = 1 - \int_0^x g(s) ds = \int_0^x \frac{|l + (c-1)(bs)^c|}{\sigma \sqrt{2\pi s^3}} e^{-\frac{(l-bs)^c}{2\sigma^2 s}} d(s) \quad (22)$$

Where the integral is normalized so that

$$\int_0^\infty g(s) ds = 1$$

Fig. 3. A. Stochastic Simulations, B. Health State and Survival and C. LE and σ



Following the above theory we have estimated the health state for the females in USA for 2010 and we have reproduced the results by applying stochastic simulations (see Figure 3A).

The health state curve for USA females in 2010 is illustrated by the heavy curve (see Figure 3B). The corresponding survival curve (continuous line) for this related case is presented with $\sigma=0.06881$. The curve with higher value for $\sigma=0.09$ is presented by a dashed line whereas the lower values for σ ($\sigma=0.05$, dotted light curve and $\sigma=0.03$ dashed upper curve) appear as well. The very important case with $\sigma=0$ corresponding to the deterministic case is illustrated with a heavy line (BDC). The latter ends at the point C at the zero health state age. All the survival curves cross around the point D at the zero health state age.

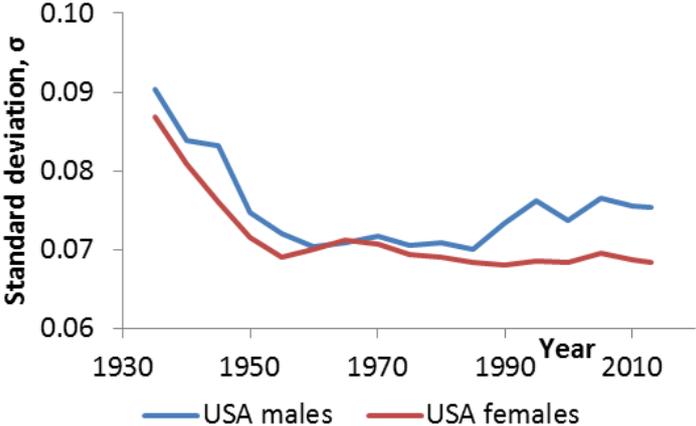
5 Comparing the Health State or Health Status and Life Expectancy

From equation (22) for the survivorship we have estimated the Life Expectancy for various values of the standard deviation σ . The results are illustrated in Figure 3C. The Life Expectancy at Birth is systematically related to the standard deviation σ . At $\sigma=0$ the Life Expectancy is equal to the age at zero Health Status (application for USA, females, 2010). The life expectancy as it is given in the life table from Human Mortality Database is 81.20 years and we estimated that this is achieved for $\sigma=0.06881$.

5.1 The time development of the standard deviation σ

As far as the standard deviation σ is estimated by the proposed method we have calculated the related figures for USA males and females and presented in Figure 4. The standard deviation and thus the dispersion is larger for males especially for the years after 1990.

Fig. 4. The time development of the standard deviation σ for USA



6 Conclusion

We have presented and applied the general theory of the health state of a population in connection to the survival curves and life expectancy. We have provided a measure for the Health State or Health Status which is independent of the dispersion parameter expressed by the standard deviation σ . We also have provided enough evidence of the relation of life expectancy and the health state or health status estimator. Finally we complete our work with stochastic simulations thus reproducing the death distribution for several countries and for different time periods.

References

Carey, J., Liedo, P., Orozco, D. and Vaupel, J. Slowing of mortality rates at older ages in large medfly cohorts. *Science* 258, 457-461 (1992)

Janssen, Jacques and Skiadas, Christos, H. Dynamic modelling of life-table data, *Applied Stochastic Models and Data Analysis*, 11, 1, 35-49 (1995).

Skiadas, Charilaos and Skiadas, Christos, H. Development, Simulation and Application of First Exit Time Densities to Life Table Data, *Communications in Statistics* 39, 2010: 444-451.

Skiadas, Christos, H. and Skiadas, Charilaos. *The Health State Function of a Population*, ISAST, Athens, (2nd Ed., January 2013) http://www.amazon.com/Health-State-Function-Population/dp/6188046505/ref=sr_1_1?s=books&ie=UTF8&qid=1364343495&sr=1-1

Skiadas, Christos, H. and Skiadas, Charilaos. *Supplement: The Health State Function of a Population*, ISAST, Athens, (December 2013).

Skiadas, Christos, H. and Skiadas, Charilaos. The First Exit Time Theory applied to Life Table Data: the Health State Function of a Population and other Characteristics, *Communications in Statistics-Theory and Methods*, 43, 2014: 1985-1600.

Skiadas, Christos, H. and Zafeiris, Konstantinos, N. Population Aging and Healthy Life: Lessons from the Related Studies, *RELIK2015* (12-13 November 2015).

Skiadas, Christos, H. and Skiadas, Charilaos. Exploring the State of a Stochastic System via Stochastic Simulations: An Interesting Inversion Problem and the Health State Function. *Methodology and Computing in Applied Probability* (2015, Volume 17, Issue 4, pp 973–982).

Torrance, George W. "Health Status Index Models: A Unified Mathematical View." *Management Science*, 22(9), 1976: 990-1001.

Weitz, J.S. and Fraser, H.B. Explaining mortality rate plateaus, *Proc. Natl. Acad. Sci. USA*, 98(26), 15383 (2001).

Contact

Christos H Skiadas

ManLab, Technical University of Crete

73100 Chania, Crete, Greece

skiadas@cmsim.net