Social inequalities in health expectancy of elderly : evidence from the HID survey

Cristina Giudici and Maria Felice Arezzo
University of Rome “LaSapienza”

1 - Introduction

The majority of analyses dealing with health, are based on cross-section surveys, and more particularly on the self- assessment of the interviewed individuals’ state of health at a given moment. Thus the results reflect the feeling and beliefs held by people concerning their health, and they are dependent on unpredictable factors linked to the life of the individual at the moment of the investigation, to recent traumatic events, and to the awareness of the individual of his/her own state of health. When the questionnaire refers to the past, the individual’s ability to remember accurately is also important. It is also necessary to underline that the calculation of health expectancy is often based on the so-called Sullivan method (1971), which apply prevalence data of incapacity (obtained through cross-sectional surveys) on age specific person years in a life table. This type of approach is limited because the incidence of incapacity in theperiod of reference is not taken into account; the prevalence observed at a given moment derives from the transitions from one state of health to another as verified in the past, and therefore depends on the history of the cohorts which make up the sample group.

A possible alternative is the method of multi-state tables, based on the analyses of the transitions between successive states of health. The information necessary for this type of analysis derive from longitudinal surveys. The result, in this case, is the so-called period (or stable) prevalence and can be interpreted analogously to the stationary population of a life table. In a life table the sequence of survivors of a fictitious cohort, subject to the laws of mortality of the moment, can be observed and could be interpreted as “the structure of an hypothetical population made up of successive cohorts of newborns, of the same initial size, which have each been reduced, age by age, by the same period death rates. […] The population is a stationary one which has reached an equilibrium structure associated with the period mortality conditions” (E. Cambois, J.M. Robine, and N. Brouard, 1999). Analogously the stable prevalence of disability can be interpreted as the proportion of the disabled amongst the survivors of successive fictitious cohorts, subject to the flows of entry on disability and recovery observed in the period under examination. The health expectancy so obtained is therefore more satisfactory than transversal measurements in order to monitor the conditions of health of a population and to forecast its future development.

The aim of this study is to estimate the health transition probabilities for elderly and to study how they are influenced by socio-relational variables, in order to calculate differences in health expectancy linked to socio-relational factors.

2. – The French survey on handicaps, disabilities and dependency (HID)

The national survey on handicaps, disabilities and dependency (HID), carried out in France, between 1998 and 2001, by the National Institute on Statistics and Economical Studies (INSEE), in collaboration with the Institut National d’Etudes Démographiques (INED), leads us to a measurement of the stable prevalence of disability. The HID survey focused on two groups: the people who live in ordinary settings (households) and the people who live in institutional settings (and particularly in what are considered as « medico-social » institutions in France). Thus, two surveys (in institutions and households) were conducted twice at a two years interval in order to estimate the disability entry and exit flows.

The sample (15,000 persons in healthcare and social institutions and 16,500 in households) is representative of the entire French population. In order to estimate the stable prevalence of disability, this study is limited to the analysis of the data regarding the population resident in private dwellings aged 55 and over (9804 interviews).

2.1 – The sample attrition

A specific problem related to the *HID-ménage* survey concerns the sample attrition. The Classification and Regression Trees (CART) allows to estimate the health status of those individuals for which no information on health evolution is available. A complete description of CART is really long and beyond the scope of this work; we will try to briefly describe it and to highlight its functioning.

CART is a supervised classification algorithm. A supervised classification problem can be summarized as follows: for *n* objects, characterised by a set of *k* features $X = (X1, X2, ..., Xi, ...Xk)$, is known a priori the class $j = 1, 2, 3...J$ they belong to.

Classes are generally indicated with variable Y. The scope is to predict which is the class a new object belong to given its characteristics. A supervised classification algorithm is a mathematical rule which assign a new object to a class *j*. A function $d(X)$, called classifier, is built in a way that it generates a partition of the feature space *X* into *J* non overlapping subsets.

CART is a binary recursive partitioning procedures capable of handling both continuous and nominal characteristics. Starting with the entire sample (parent node), it divides it into two children nodes; any of them are then divided into two grandchildren. A node is said to be final if it cannot be divided. The procedure stops when the tree is grown at its maximum size. The full grown tree is then pruned back in order to look for the best final tree. This is the one that minimize the so called cost-complexity function which is a function that takes into account at the same time the misclassification rate of individuals and the total number of final nodes.

The original data has a certain level of heterogeneity: if all individuals belong to the same class, there is no heterogeneity in the data. On the opposite side if individuals are uniformly distributed among the *J* classes the heterogeneity reaches its maximum level. Any split is done according to a variable *Xi*: the algorithm searches over all feature space looking for the optimal division that is for the binary split that reduces data heterogeneity most. Impurity reduction can be measured and it gives variables ranking based on their capability to separate object. This is called variable importance.

An important issue is the capability of a tree to correctly classify a new individual. A measure of this generalization power is the misclassification rate which is simply the number of misclassified individuals out of all observed individuals. If the original sample is big enough, a good estimate of the true misclassification rate is obtained by randomly splitting the sample in two sub samples and use the first part of the data (normally 70% of it) to grow the tree and the second to test it.

3 – The analysis of health transition probabilities

On the basis of the HID survey, health is measured through a functional approach, in particular the survey allows the measurement of possible alterations of the physiological functions of the organism and restrictions in the carrying out of everyday tasks.

The study refers to three indicators: physical, mental and self-perceived health status.

From a physical point of view the study refers to the so called Activities of Daily Living (ADL) i.e. getting washed, getting dressed, eating pre-cooked foods, going to bathroom, getting out of the bed or up of a chair. Concerning mental health variables refer to the temporal and spatial orientation: if individuals can remember what time of the day it is and if they can find their way home.

For each of these questions the possible answers are grouped in four categories: no problem; some difficulties; great difficulties; help needed. We adopted a large definition of disability, considering disabled those individuals which have any difficulties or need help in their daily activity (by the same token mental disability refers to a frequent or permanent lack of orientation in at least one of the two considered domains).

The logistic analysis shows the importance of the social and relational context in worsening the physical, mental and self-perceived health status. In particular, results highlights the role of the social and physical isolation; in particular, the existence and availability of social and family contacts and the information on whether or not individuals have lived through a dramatic experience (usually the death of a relative) in the last two years have been considered. The probability to move from disability free to a disabled situation is influenced by the general health status: in particular physical health influence self-perceived health transition and *vice-versa*, but the analysis shows also the effect of other covariates: generally, the existence of a family and social environment with which individual interact seems to be a strong factor in protecting health, except the number of children which shows a slight opposite effect. The involvement in cultural activities have also a positive influence on health.

Table 1 - Logistic regression on worsening health status (Activity of daily living)

constant	-5,81	306681,785
Self perceived health		
medium	ref	
good	-0,42	26162,719
poor	0,41	9132,870
Living far from public services, alimentary and medicine shops etc.		
yes	1,42	27638,773
no	ref	
age	0,04	
Involvement in cultural activities		
yes	ref	
no	0,44	28761,658
nombre d'enfants	0,10	23904,381
Having a moral support		
no, don't need any moral support	1,07	116832,454
no, but moral support needed	0,36	1561,835
yes	ref	
Social position		
administrative	0,13	1648,259
artisan, shop or business owner	-0,21	2280,339
farmer	-0,19	2026,478
intermediate occupation	0,08	342,884
other inactive	-0,51	4823,474
manager, higher occupation	0,23	2932,136
manual labourer	ref	

Table 2 - Logistic regression on worsening health status (Mental health)

Costant	-1,942	38886,642
Functional limitation (ability to fill in a form)		
no	ref	
yes	,438	17301,165
Functional limitation (hearing, sight, speech)		
yes	ref	
no	-,272	15112,080
Self perceived health		
medium	ref	
good	-,716	121125,849
poor	,547	26253,498
Age	,020	23630,486
Participation in one ore more associations		
no	ref	
yes	-,145	4984,537
Involvement in cultural activities		
yes	ref	
no	,195	8240,137
Number of kids	,077	16789,403
Number of brothers and sisters	-,051	11783,462
Having a moral support		
no, don't need any moral support	ref	
no, but moral support needed	,226	6070,759
yes	-,333	1419,414
Marital status		
married	ref	
divorced	-,167	1080,184
separed	1,039	10995,650
single	,724	41933,564
widowed	-,408	25024,294
Social position		73759,481
administrative	-,377	19793,017
artisan, shop or business owner	-,571	27882,089
farmer	-,603	32860,719
intermediate occupation	-,467	19021,135
other inactive	-1,191	39493,773
manager, higher occupation	-,356	8170,262
manual labourer	ref	
Dramatic experience lived through in the last two years		
no	ref	
yes	,458	59855,931

Table 3 - Logistic regression on worsening health status (Self perceived health)

Costant	-1,502	129,189
Functional limitation (ADL)		
no	-,247	4,687
yes	ref	
Living far from public services, alimentary and medicine shops etc.		
yes	,882	10,825
no	ref	
Involvement in cultural activities		
yes	ref	
no	,292	11,607
Sex		
men	//	//
women	ref	
Functional limitation (moving inside the house)		
yes	,667	54,759
no	ref	

4 – The measurement of life expectancy without disabilities

On the basis of the perceived ability to carry the mentioned activities, it is possible to construct the age-specific flows of entry into and exit from disability, and the matrix of the transitions between good health (1), disability (2) and death (3) These three states make up the elements of a non-homogeneous Markov chain, in which the probabilities of transition depend on age and, if necessary, on other covariates; and the states of transition are discrete.
The probability for an individual aged x, observed in the state i during the first transition, to find him/herself in state j during the second transition is indicated with ${}_xp_{ij}$. Every probability of transition can be expressed using the following logistic multinomial relationship:

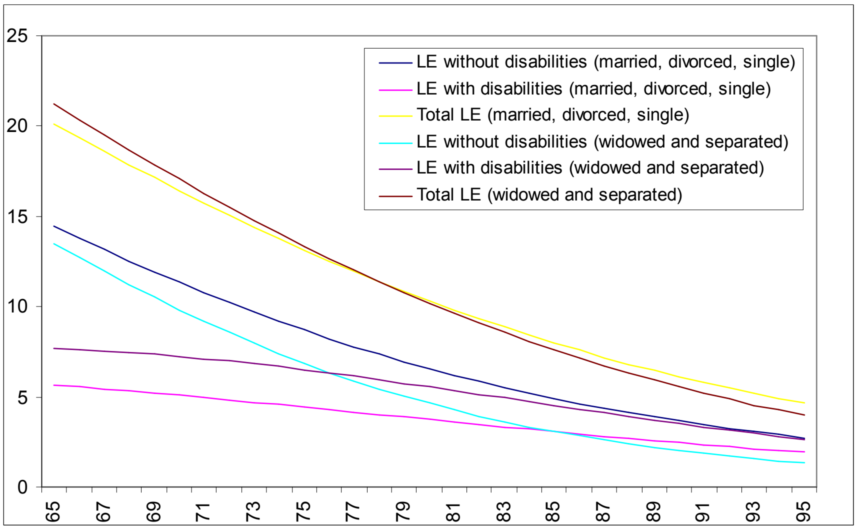
[1] $\ln({}_xp_{ij} / {}_xp_{ii}) = a_{ij} + b_{ij} * x$

in which p_{ij} indicates the probability of transition from state i to state j and x indicates the age.
The non-absorbing states are defined respectively in terms of physical and mental disabilities and bad self perceived health, while death is the absorbing state.
The simulation of a cohort of births which experience over time the observed transitions of health, leads us to the calculation of the health expectancy for people respectively with and without disabilities at the initial age (55 years)[1].
Main results confirm that the relational environment is an important factor in protecting health; for example, figure 2 shows the life expectancy for individuals in good health at initial age and its decomposition into health expectancies, by marital status: people widowed and separated show shorter life expectancy without physical disabilities: at 65 years old the married, divorced and single people can expect to live on average 14,5 years without disability, 5,5 with disabilities, and 20 years of total life expectancy. At the same age, widowed and separated people can expect to live on average one year more of total life expectancy, but two year less without disability. The introduction into the analyses of other demographic and social information confirms the initial

[1] The analyses was carried out with the help of the program IMaCh (Interpolation of Markov Chains), developed by Agnèse Lièvre e Nicolas Brouard, and available on the website of the Institut National d'Etudes Demographiques (INED), France.

hypothesis. In particular, main results underlines the importance of the moral support: people declaring to have a moral support show an higher total life expectancy, due to a longer disability free life expectancy. At 65 years old the gap is greater than one year and it diminishes slowly with age.

Figure 1 – Life expectancy for individuals in good health at initial age and its decomposition into health expectancies (adl) , by marital status



5 - Conclusions

Contemporary European societies live better and longer in comparison to the past. That is to say that alongside demographic aging there is also an undeniable biological rejuvenation. In this context the measurement of life expectancy in good health, which in some European populations reaches close to 75 years, recalls on the one hand the hypotheses of lengthening of active life outlined in the international debate. On the other hand, it gives the opportunity to return the entire life cycle to the debate, rethinking the definitions and measurements of aging and transplanting a wide band of the population from the category of the elderly to that of adults, in a context of flexibility and availability of the opportunity to move in and out of economic activity.

6 – References

Breiman, L, Friedman J.H., Olshen R.A., Stone C.J. (1984). *Classification and Regression Trees.* Wadsworth, Belmont, CA.

Cagiano de Azevedo R.(ed.), The European Welfare in a Counteraging Society, ed. Kappa, Roma, 2004

Cambois, J.M. Robine, and N. Brouard, 1999, “Life Expectancies Applied to Specific Statuses. A History of the Indicators and Methods of Calculation”, in Population, an English Selection, vol. 11 (1999), pp. 7-34

Giudici C., « Les déterminants socio-démographiques de la santé aux grands âges » working paper, Les Lundis de l’INED, Parigi, 22 maggio 2006

Hastie T., Tibshirani R., Friedman, J. (2001). *The Elements of Statistical Learning.* Springer.

Laditka J.N. and S.B. Laditka, “Increased hospitalization risk for recently widowed older women and protective effects of social contacts”, Journal of women and aging, vol. 15 n°2/3 2003

Robine J.M., Jagger C., “Allongement de la vie et état de santé de la population”, Population et Société, Volume VI, Editions de l’INED, 2004, p. 51-84

Robine J.M., C. Jagger, C.D. Mathers, E.M. Crimmins, R.M. Suzman, *Determining Health Expectancies*, ed. Wiley, England 2003

Sermet C., E. Cambois, « Measuring the State of Health », in G. Caselli, J. Vallin, G. Wunsch, Demography, Analysis and Synthesis, vol. 2, Academic Press Elsevier, 2006

Wu, X. Kumar, V. (2009). *The Top Ten Algorithms in Data Mining.* Chapman & Hall/CRC.