Biometrical Letters
Vol. 43 (2006), No. 1, 1-8

Permanent disability and social security

Inês Jorge Sequeira, João Tiago Mexia

Department of Mathematics, Faculty of Science and Technology, New University of Lisbon, Quinta da Torre 2829-516 Caparica, Portugal
Tel.: +351 21 294 83 88, Fax: +351 21 294 83 91
e-mail: ijs@fct.unl.pt

SUMMARY

A Markov chain approach is used to estimate the probability of a man covered by Portuguese Social Security becoming permanently disabled. This estimator is used to assess the contribution of each active man covered by the Portuguese Social Security to insure against that possibility. This prediction is necessary due to the pay-as-you-go system used.

Key words: Pay-as-you-go, permanently disabled, prediction of yearly contributions

1. Introduction

This paper attempts to be a contribution to creating a sounder pay-as-you-go policy for Social Security.

Since under this system active members have to pay for the pensions received by retired persons, it is extremely important to predict the amounts to be paid.

Assuming that, as in Portugal, people permanently disabled are entitled to a pension, the corresponding amount must be predicted. For safety, we will also require prediction bounds. A practical way of formulating such amounts and bounds is to obtain the corresponding value per year for each active contributor.

To achieve this we used Markov chains, for instance see Parzen (1965). The transition probabilities, for each sex, will be

- q_x - probability of an active contributor dying at age x;
- q'_x - probability of a permanently disabled person dying at age x;
- i_x - probability of an active contributor becoming permanently disabled at age x

Besides these in our model we considered entry probabilities for ages between 20 and 30.

Now probabilities q_x may be obtained from life tables and suitably modified to give probabilities q'_x (we increased the corresponding probabilities by 10%).
Moreover there is usually data available from which to estimate the entry probabilities.
Thus the main problem in building our model was obtaining the probabilities i_x. This problem will be discussed in the next section. In the third section we show how to compute the total probability of being permanently disabled, and in the fourth section we carry out the prediction of the cost of permanent disability.
We consider the male Portuguese population; for the female population we could proceed in the same way.

2. Permanent disability

2.1. Estimation

Assuming that active persons become permanently disabled independently of each other and that there are R_x active persons (of a given sex) of age x, the number S_x of persons that get permanently disabled at age x will have binomial distribution with parameters R_x and i_x, thus we get the estimator

$$\hat{i}_x = \frac{S_x}{R_x}$$

We applied this estimator to a set of Portuguese data.

2.2. Disability strength

Given a population with initial size l_0 that decreases according to the differential equation:

$$\frac{\frac{\partial l_x}{\partial x}}{l_x} = -\mu_x$$

integrating we get $l_x = l_0 \exp(-\int_0^x \mu_t dt)$ at age x.

Thus the probability of a person belonging at age $x+1$ given that he belongs at age x will be,

$$r_x = \frac{l_{x+1}}{l_x}$$

This kind of model has been used to obtain mortality laws such as

Gompertz (1825): $\mu_x = Ce^{\delta x}$

Makeham (1860): $\mu_x = A + Ce^{\delta x}$

Bernardino (1994): $\mu_x = A + Bx + Ce^{\delta x}$

from which mortality tables can be built. Actually we will use tables based on the last of these laws, which is well-suited to the Portuguese population, see Bernardino (1994).

We also use the last law for permanent disability. Since

$$\int_0^x \mu_t dt = Ax + \frac{B}{2}x^2 + \frac{C}{\delta}(e^{\delta x} - 1)$$

we get

$$l_x = l_0 e^{-Ax - \frac{B}{2}x^2 - \frac{C}{\delta}(e^{\delta x} - 1)}$$

Taking

$$\log(s) = -A\text{, } \log(r) = -\frac{B}{2}\text{, } \log(g) = -\frac{C}{\delta} \text{ and } h = e^{\delta}$$

we have

$$l_x = l_0 s^x r^{x^2} g^{(h^x - 1)}$$

Thus we have

$$r_x = s r^{2x+1} g^{h^x (h-1)}$$

so that

$$\log(r_x) = \log(s) + (2x+1)\log(r) + h^x (h-1)\log(g)$$

Putting

$$\log(s) + \log(r) = -\alpha\text{, } \log r^2 = -\beta \text{ and } (h-1)\log g = -\gamma$$

we get

$$\log(r_x) = -\alpha - \beta x - \gamma e^{\delta x}$$

We now look at r_x as the probability of not becoming permanently disabled at age x, then

$$i_x = 1 - \exp(-\alpha - \beta x - \gamma e^{\delta x}) = 1 - r_x$$

2.3. Semi-linearization

The expression for i_x is not linear, but if we assume δ known, taking $x_1 = x$ and $x_2 = e^{\delta x}$ we have

$$z = -\log(1 - i_x) = \alpha + \beta x_1 + \gamma x_2 \quad (1)$$

Given the pair (x, i_x) we get the corresponding triplet (x_1, x_2, z) and carry out a least-squares adjustment. For each δ we have a sum $s(\delta)$ of squares of residues. We observe that minimizing $s(\delta)$ is equivalent to maximizing

$$r^2(\delta) = 1 - \frac{s(\delta)}{\sum_{k=1}^{n} (\hat{i}_k - \bar{\hat{i}})^2}$$

So we are led to choose δ through the minimization of $s(\delta)$. Usually this is done numerically. We obtain the $s(\delta_0 + j\Delta_1)$, $j = 1, \ldots, m$ with $[\delta_0, \delta_0 + m\Delta_1]$ a range such that we are confident that it contains δ. From the results in this first iteration we choose a sub-range $[\delta_0 + j'\Delta_1, \delta_0 + j''\Delta_1]$ containing the lesser $s(\delta_0 + j\Delta_1)$. We now consider a finer partition (with $\Delta_2 < \Delta_1$) for this sub-range and repeat the procedure until the differences between the minimum $s(\delta_0 + j\Delta_m)$ and the adjacent values are sufficiently small.
Actually the number of iterations required for our data, on the male population, was 3. The results are shown in Figures 1, 2 and 3.

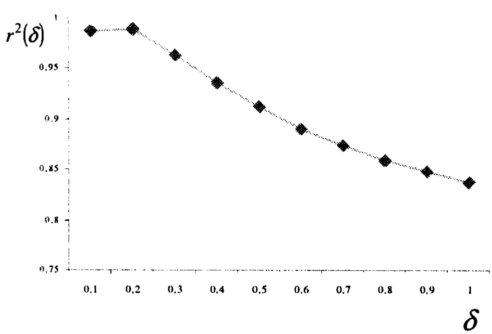

Fig. 1: δ between 0 and 1 with increment 0.1

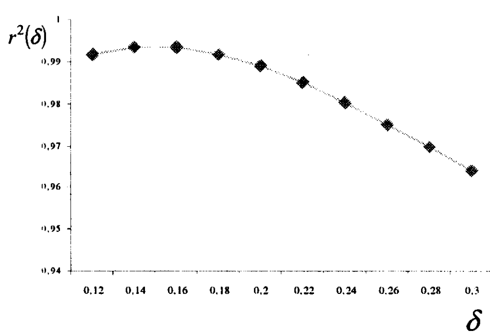

Fig. 2: δ between 0.1 and 0.3 with increment 0.02

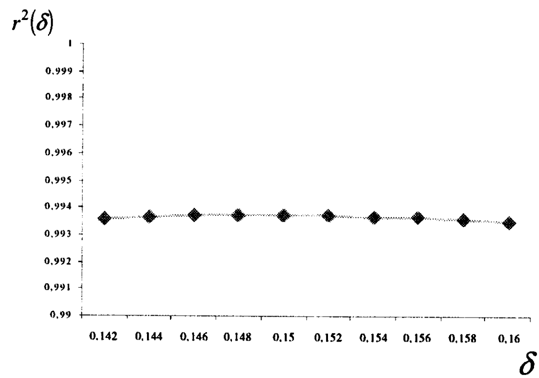

Fig. 3: δ between 0.14 and 0.16 with increment 0.002

Thus, with $\widehat{\delta}=0.15$ we obtained $r^2(\delta)=0.9937$, which can be considered an excellent fit. With $\widehat{\delta}$ the chosen value for δ, we take as estimators for α, β and γ those obtained for the linear regression corresponding to $\widehat{\delta}$. The results are presented in Table 1.

Table 1: Estimators for parameters

$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\gamma}$
0.000380561	-3.48376E-05	9.17389E-05

Finally Figure 4 illustrates the probabilities i_x for ages between 20 and 65.

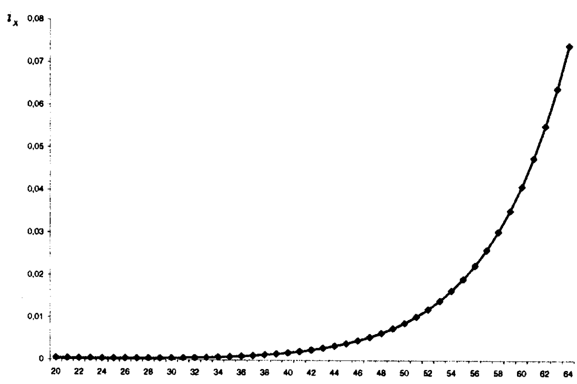

Fig. 4: Evolution of probabilities of permanent disability

The values for these probabilities are presented in the appendix.

3. Total probability of disability

The probability i of a randomly chosen man or woman being permanently disabled will be

$$i = p_{20}\frac{\sum_{x=20}^{64} j_x s_x k_x}{\sum_{x=20}^{64} s_x} + p_{21}\frac{\sum_{x=21}^{64} j_x s_x k_x}{\sum_{x=21}^{64} s_x} + \cdots + p_{30}\frac{\sum_{x=30}^{64} j_x s_x k_x}{\sum_{x=30}^{64} s_x}$$

so we have only to obtain the values of the probabilities involved.
Starting with the entry probabilities p_x we can compute these from the numbers of admissions, given by $R_x - R_{x-1}(1-q_{x-1}-i_{x-1})$, $x = 20,\ldots,30$ where R_x is the number of active men of age x.
To use this last expression we use the values of i_x obtained in the previous section and take those of q_x from the Portuguese mortality table PMP/94, which is presented in the appendix, see Mexia & Côrte Real (1995). The values obtained for the p_x are presented in Table 2.

Table 2: Probabilities per age of entrance

Age	20	21	22	23	24	25	26
Prob.	0.624861	0.078963	0.063966	0.051011	0.042801	0.0333	0.030411

Age	27	28	29	30	31	...
Prob.	0.029604	0.019862	0.017537	0.007687	0	...

Now the s_x; $x = 20,\ldots,64$ will be the probabilities of people of age x being active. To obtain these probabilities we take $s_{20} = 1$ and $s_x = s_{x-1}(1-q_{x-1}-i_{x-1})$; $x = 21,\ldots,64$.
Next, the j_x will be the probabilities of an healthy person at age $x-1$ becoming permanently disabled at age x. We can take $j_{20} = i_{20}$ and $j_x = s_{x-1}i_x$; $x = 21,\ldots,64$.
Lastly, k_x will be the probability of a disabled person with age x arriving at the final age of 65. Since we increased by 10% the mortality rates for disabled persons, one of these at age x will have the probability $1-1.1q_x$ of attaining age $x+1$, probability $(1-1.1q_x)(1-1.1q_{x+1})$ of attaining age $x+2$ and probability $k_x = \prod_{l=x}^{64}(1-1.1q_l)$ of attaining the retirement age of 65. We thus obtained the estimator $\hat{i} = 0.005964$.

4. Prediction

Let us assume that we have a stable population of size P. Portuguese Social Security statistics give the value $P = 4845662$.
Since the total probability of disability i is small we may assume that the number N of permanently disabled persons has a Poisson distribution with parameter iP.
Since N is large we may also assume – see Guerreiro, Mexia & Sequeira (2005) – that

$$Z = \frac{N - iP}{\sqrt{iP}} \sim N(0.1)$$

Thus for a minimum pension of $c = €3038$ per year we get the 95% confidence range (in euros) of $[17.83;18.41]$ for the individual yearly contributions from active members.

Appendix

Age interval (in years) x to x+1	Probability of dying in interval [x,x+1] 1000qx	Probability of dying increased 10% in interval [x,x+1] 1100qx	Probability of becoming permanently disabled at age x ix
20	1.75	1.925	0.00045
21	1.937	2.1307	0.00043
22	1.742	1.9162	0.00042
23	1.953	2.1483	0.00041
24	2.217	2.4387	0.00040
25	2.441	2.6851	0.00040
26	2.252	2.4772	0.00040
27	2.248	2.4728	0.00041
28	1.991	2.1901	0.00042
29	2.329	2.5619	0.00044
30	2.554	2.8094	0.00047
31	2.884	3.1724	0.00052
32	2.719	2.9909	0.00057
33	2.672	2.9392	0.00064
34	2.985	3.2835	0.00073
35	2.772	3.0492	0.00083
36	3.122	3.4342	0.00096

37	3.172	3.4892	0.00112
38	3.439	3.7829	0.00130
39	3.629	3.9919	0.00153
40	3.522	3.8742	0.00179
41	3.429	3.7719	0.00210
42	3.538	3.8918	0.00247
43	4.216	4.6376	0.00290
44	3.335	3.6685	0.00340
45	4.471	4.9181	0.00400
46	4.744	5.2184	0.00469
47	5.205	5.7255	0.00551
48	5.321	5.8531	0.00646
49	4.99	5.489	0.00756
50	5.838	6.4218	0.00886
51	6.235	6.8585	0.01036
52	7.679	8.4469	0.01211
53	8.479	9.3269	0.01414
54	9.344	10.2784	0.01651
55	10.525	11.5775	0.01925
56	11.133	12.2463	0.02243
57	11.733	12.9063	0.02613
58	12.504	13.7544	0.03040
59	14.079	15.4869	0.03536
60	16.221	17.8431	0.04108
61	16.239	17.8629	0.04770
62	16.894	18.5834	0.05534
63	18.742	20.6162	0.06414
64	21.472	23.6192	0.07426

References

Bernardino, G. (1994), Ajustamento Analítico de Taxas de Mortalidade, Master's Thesis, Faculty of Science and Technology, New University of Lisbon. (In Portuguese)

Guerreiro, G.R., Mexia, J.T. and Sequeira, I.J. (2005), Invited presentation, 25th International Symposium on Forecasting, San Antonio, Texas, June 12-15

Mexia, J.T. and Côrte Real, P. (1995), Tabelas de Mortalidade Auto-Correctivas: O Caso Português, *Boletim do Instituto dos Actuários Portugueses*, no. 35, 1-120

Parzen, E. (1965), Stochastic Processes, *Holden Day,* San Francisco

Sequeira, I.J. (2002), Tabelas de Invalidez, Master's Thesis, *Instituto Superior de Economia e Gestão.* (In Portuguese)