Actuarial Analysis via Branching Processes

Carlos Alberto de Bragança Pereira, ***cpereira@ime.usp.br*****, Statistics Department,**
Fabio Nakano, ***nakano@supremum.com*****, Supremum Assessoria e Consultoria, and**
Julio Michael Stern, ***jstern@ime.usp.br*****, Computer Science Department,**

Institute of Mathematics and Statistics of the University of São Paulo

ABSTRACT

We describe a software system for the analysis of defined benefit actuarial plans. The system uses a recursive formulation of the actuarial stochastic processes to implement precise and efficient computations of individual and group cash flows.

Keywords: Actuarial modeling, branching processes, defined benefit pension plans, recursive functions.

1. INTRODUCTION

We report the use of a software tool for the analysis of cash flows due to pension plans (PP) in Brazil. Many of the existing pension funds are of defined benefits (DB) type, where the retired member or his surviving dependents receive a lifelong monthly income. The subjacent stochastic process is modeled as a branching process driven by several time dependent hazard rates. The expected cash flows are computed by recursive functions describing the branching process, so avoiding several approximations used in standard actuarial methods. These recursive functions also give a direct calculation of the cash flow's variance and other statistics.

2. THE BASIC MODEL

The main benefit for a DB PP (defined benefit pension plan) member is a lifelong retirement monthly income. Prior to his retirement a member is named active. The retirement income is a function of the active member's past incomes or contributions (ex. last periods average). The active member makes contributions to the pension plan, and these contributions can be complemented by contributions from a sponsor (ex. employer or government). An active member will become inactive when retired, at a maturity time, or earlier if disabled (ex. injury or disease). An active member can also withdraw from the PP.

The member may have dependents (usually his family) entitled to a pension monthly income after the member's death. Dependents may be permanent, who will receive a lifelong pension (ex. wife/widow, disabled children), or temporary, who will receive the pension for a limited time (ex. normal children up to maturity age of 21). Each dependent's pension is a fraction of the member's retirement income. An additional one time (lump sum) death assistance may also be available to the family.

Several constraints and corrections [1] [4] [5] [6] [7] [17] increase the complexity of this basic model, for example:

- The retirement, and all other benefits defined by it, may be corrected by a long term inflation index, or may be adjusted by the income of an active member of the same status of the retired one.

- The retirement maturity time may be based the member's age and employment time, and also on the PP rules and government regulations, both changing over time.

- The members may receive a basic government retirement, being the PP obligation to supplement it up to the PP's DBs.

- Changing social habits and legal definitions may change the status of entitled dependents (ex. mistresses and out of wedlock children).

- Withdrawing members may claim his (or also the sponsor's) contributions corrected by inflation or financial investment indices.

3. Graphs and recursive Formulation

A branching process is described by a graph, where each vertex (or node) corresponds to a state, and each arch (or edge) connecting two vertices corresponds to a possible state transition. In the actuarial processes we are studying, a state is characterized by the member's age, time of employment, salary, family, etc. A transition is characterized by its probability, as well as by the benefits and contributions the transition implies. Usually it is convenient to give the benefits and contributions values as fractions of the main benefit (retirement), or some other adimensional unit.

The expected value of a member's random variable (ex. benefits or contributions) at a given period, is its probability weighted sum of the random variable's value at all possible transitions at that period: *E(X(t)) = sum(j in W(t)) Pr(j)*x(j)*, where *W* is the set of all possible transitions, *x(j)* the random variable value at that transition, and *Pr(j)* the transition's probability. That random variable expected (cash) flow is the array of its expected values in the future (subsequent periods, usually years). The graph description of the branching processes gives a recursive algorithmic formulation for the computation of all these cash flows.

4. Retired Member Graph

A retired member state has its age, benefits, and list of dependents. Let us assume that a retired member has at most one permanent dependent (wife). If the member and his wife are both alive at time *t*, the member will be, at time *t+1*, in one of four possible states, depending on his and his wife survival or not: Let the retired member's and his wife's ages be *(x, y)* at time *t*. He can reach at time *t+1* the states *(x+1, y+1), (x+1,~), (~, y+1), (~, ~)*, where the tilde *(~)* means death. The probability of each of the four transitions are given by the force of mortality, *h(a)*, at the respective ages:

Pr(t, (x, y), (x+1, y+1)) = (1-h(x))(1-h(y));* Eq. (1)
*Pr(t, (x, y), (x+1, ~)) = (1-h(x))*h(y) ;* Eq. (2)
Pr(t, (x, y), (~, y+1)) = h(x)(1-h(y)) ;* Eq. (3)
*Pr(t, (x, y), (~, ~)) = h(x)*h(y) ;* Eq. (4)

A retired member leaves the system (PP) when all cash flows by him generated cease to exist, possibly long after his own death. The leaves of the retired member branching tree are the terminal state *(~,~)*. Temporary dependents (children) are supposed to always (deterministically) survive up to maturity age.

As we have mentioned in section 2, multiple permanent dependents may occur. One possibility would be to incorporate the multiple permanent dependents directly in the branching process, at a heavy computational cost. It so happens that the standard pension rules of DB PPs only take into account the total number of dependent survivors after the members death. This allows a significant simplification: We model the permanent dependent in the retirement branching process as a virtual permanent dependent corresponding to the last surviving real permanent dependent. In appendix 1 we list a small Matlab program to compute the cumulative life probability distribution of such a virtual dependent. It is easy to generalize the procedure to three or more permanent dependents. The cash flows of permanent dependents deceasing earlier than the last survivor can then be modeled as independent cash flows.

The precise modeling of the multiple permanent dependents effect has a significant impact on those members' benefit's expected cash flows (typically 30%). Since this situation is increasingly more frequent, such careful analysis is recommended. Figures 1 to 4 show comparative life distributions as computed in appendix 1. Sometimes the last order statistic is approximated by the survival rates of the youngest permanent dependent. From figures 2 to 4 can see that this approximation can be quite misleading.

5. Active Member Graph

An active member sate has its age, time of membership, time of employment, education, salary, etc. While active, it is hard to obtain a reliable list of dependents, so active members are assumed to have a <u>standard family</u>, based on statistical data and the member's general profile. If a member is active at time *t*, with age *a* and employment time *e*, he will reach at time *t+1* one of four possible states, depending on he still being in the PP, active, alive, and able. Death, disability, and withdrawal are competing risks, with hazard functions (conditional on the non occurrence of the preceding risks) *hd(a), hb(a)* and *hw(e)*. So the transition probabilities (except for deterministic retirement at maturity) for death, disability, withdrawal, and remaining active are, respectively:

hd(a), hb(a), hw(e), and Eq. (5)
(1-hd(a))(1-hb(a))*(1-hw(e))*. Eq. (6)

If the member withdraws he receives a lump sum based on his past contributions. If he dies or becomes disabled, he prematurely (in comparison to maturity) enters retirement. The active member branching process is therefore limited to the main stem of surviving all risks, a structure resembling a “bamboo” more than a “tree”. The bamboo leaves are the terminal withdrawal state, or the root of a retirement branching process.

6. LIFE TABLES AND OTHER ADJUSTMENTS

Life tables: Force of mortality tables are available for several countries. The most commonly used table in Brazil is EB-7. However, a specific population, like the members of a given company or PP, can significantly depart from national averages. For specific PPs, some with up to two hundred thousand members, we had the need to adjust these tables. Figures 5 to 8 give some comparisons of these survival distributions. As usual in actuarial sciences we establish a cut-off, limiting individual age to a maximum (ex. 100 years). The impact of these adjustments on the PP total liability is considerable, up to 20%.

We used a polynomial GMDH model (Group Method Data Handling) using the available tables (prior information) and the PP population historic (observed and censored deaths) [10]. The GMDH polynomial models have variable complexity and several parameters. The best model was automatically selected by an heuristic search controlled by the PSE criterion (Predicted Squared Error) [2]. The PSE criterion’s objective is to minimize errors on yet unobserved data, compromising training data error and an overfit penalty. The final model was validated using computer intensive statistical resampling methods [12] [21].

Fractional Age Correction: While modeling a transition between consecutive periods, from t to $t+1$ (depending on how de model is implemented) unrealistic assumptions may be introduced, for example: A death transition may imply that the member dies at the very first (or very last) month of the year. To correct such a boolean (0-1) dichotomy, we may assume that the death occurs at the middle month, and use a correction factor $6/12 = 1/2$, or that the death occurs at the middle day of the middle month, and use a correction factor $(6+1/2)/12 = 13/24$, and so on. These correction factors are called fractional corrections (or discretization corrections) [5]. Their impact on the final calculations is usually small, but they are important to preserve model consistency.

Income Growth: An active member income (or salary), the basis for his benefits, is supposed to evolve with his professional life. The income usually increases over time, but such increase has a saturation effect. Several models adjust well to this situation [18], like the Modified exponential, Gompertz and Logistic (Pearl) models:

$$M(t) = a - b*exp(-c*t) ; \quad \text{Eq. (7)}$$
$$G(t) = exp(a - b*exp(-c*t)) ; \quad \text{Eq. (8)}$$
$$L(t) = a / (1 + b*exp(-c*t)) ; \quad \text{Eq. (9)}$$

7. IMPLEMENTATION

The calculation engine was implemented in plain ANSI-C programming language, in order to obtain a carefully optimized code. Intermediate lookup tables considerably speed up the computation of a PP many members cash flows. A PP with a population of 100.000 members takes about 3 hours of processing time on a Pentium 750MHz machine (MSWindows or Linux).

A GUI (Graphical User Interface), written in Delphi, provides an intuitive and easy to customize interface to the corporate user. A Delphi multi-platform data transfer interface downloads and updates the necessary data on a local database (ex. AWK or Access) from the corporate environment (ex. DB2 on an IBM-AS-400).

The analysis and simulations made with the actuarial system are used as inputs to the PP’s financial portfolio management. Several optimization models, usually employing dynamic and stochastic programming, are used with this objective [3] [9] [13] [14] [15] [16] [20] [22].

8. ACKNOWLEDGMENTS

The GUI and multi-platform data transfer modules were written at RAM Computer Systems, by Rodrigo Cascão, Augusto Andraus and Michel Cusnir. The project integrator at the first PP to be analyzed was UniSoma – Mathematica para a Produtividade. We thank Prof. Miguel Taube neto, UniSoma’s President, for his enduring patience, firm determination, and practical wisdom when things got rough. Marcos Cesar de Castro Lima and Ricardo Chagas Cruz, from Quality Corretora de Mercadorias Ltda helped us as integrators at other customers.

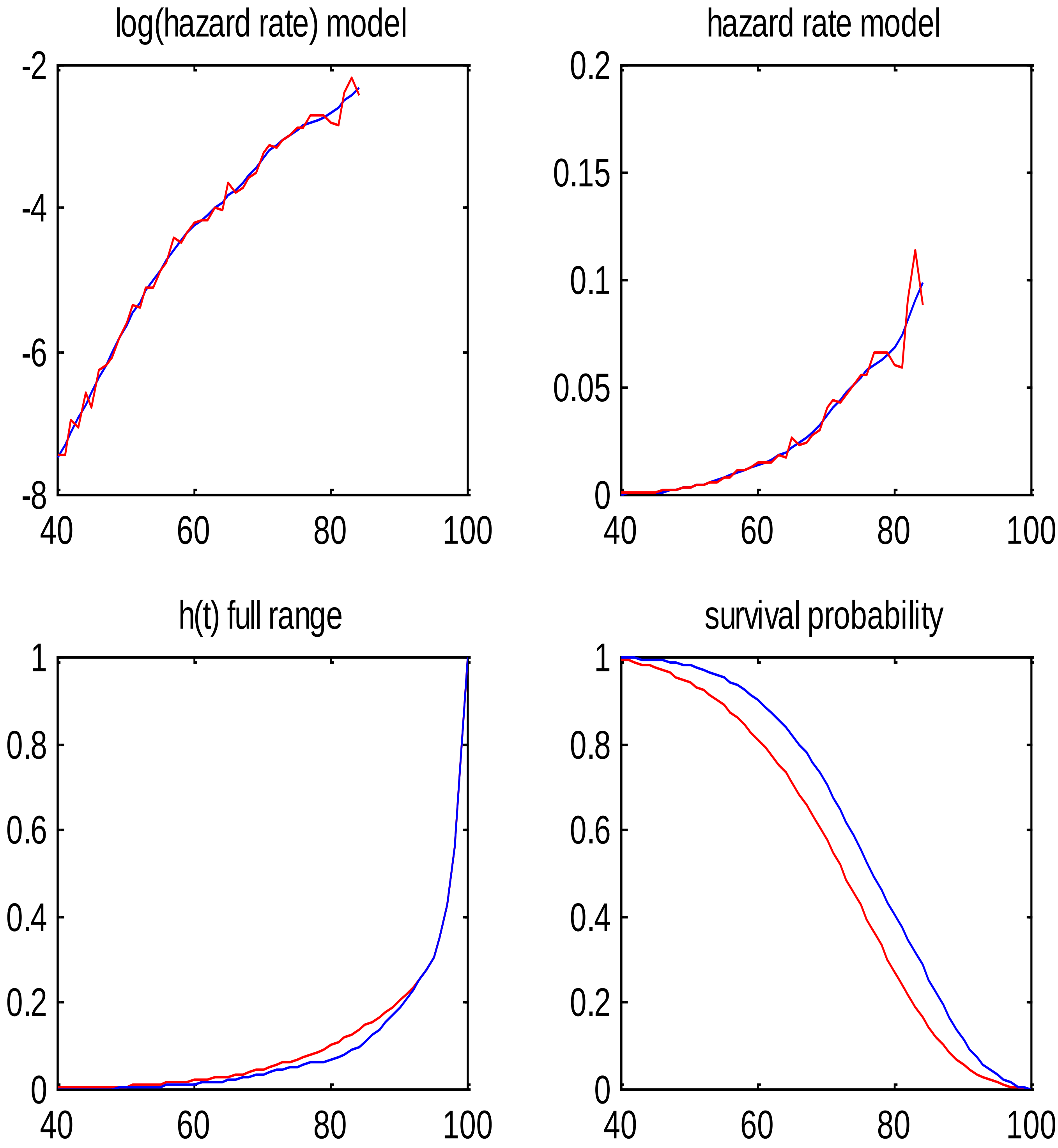
References

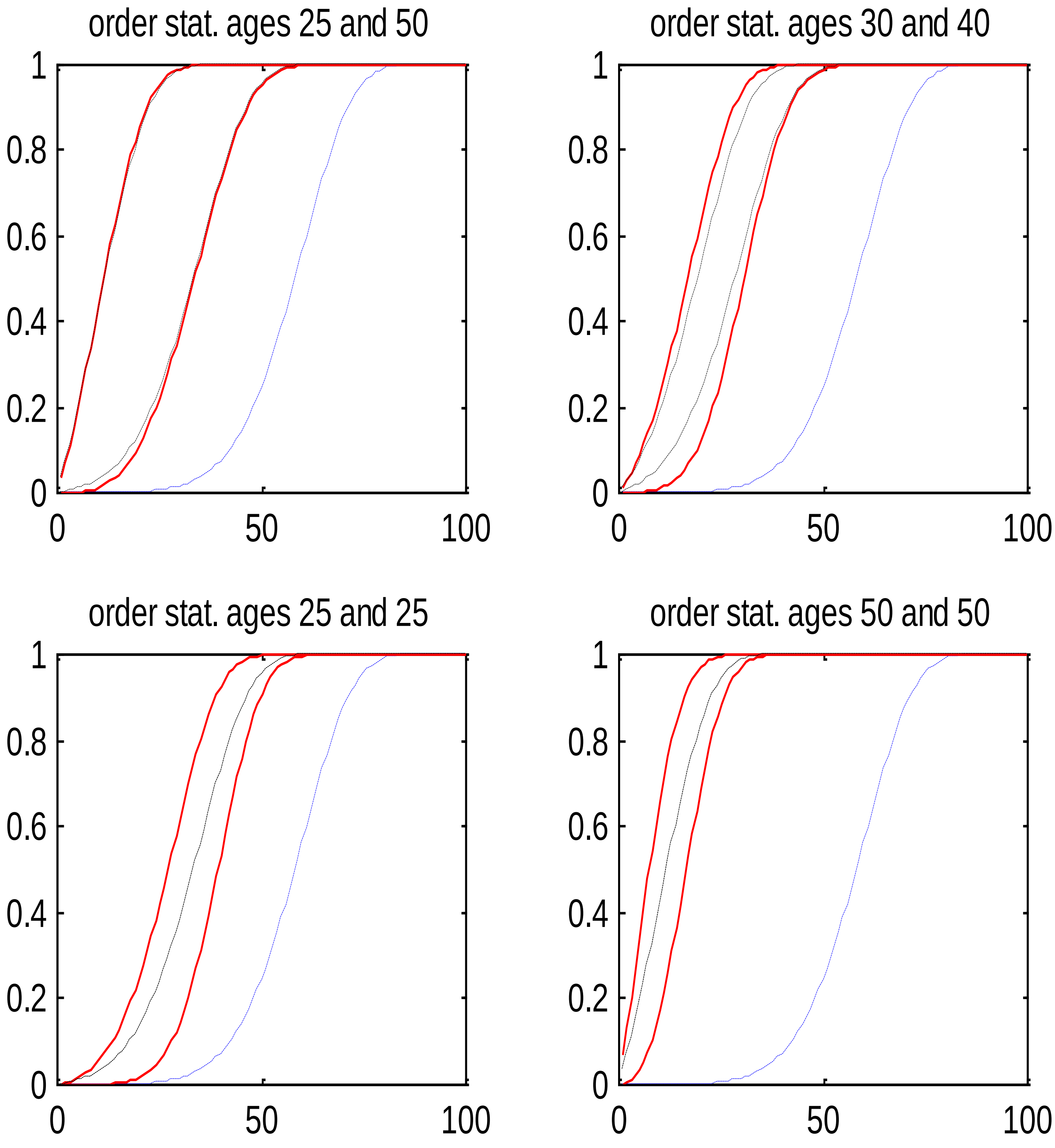
[1] U.N.de Alba J.V.Asensio. Matemática Actuarial. Instituto de Ciencias del Seguro, Editorial MAPFRE, Madrid, 1993.
[2] A..R.Barron. Predicted Squared Error: A Criterion for Automatic Model Selection, in [10].
[3] J.R.Birge F.Louveaux. Introduction to Stochastic Programming. Springer, 1997.

[4] K.H.Borch. Economics of Insurance. North-Holland, 1992.
[5] N.L.Bowers Jr. H.U.Gerber J.C.Hickman D.A.Jones C.J.Nesbitt. Actuarial Mathematics. The Society of Actuaries, USA, 1997.
[6] C.D.Daykin T.Pentikäinen M.Pesonen. Practical Risk Theory for Actuaries. Chapman & Hall, 1994.
[7] B.de Finetti. Scritti. Pitagora Editrice, Boligna, 1991.
[8] M. DeGroot. Probability and Statistics. Wesley 1986.
[9] E.V.Denardo. Dynamic Programming. Prentice-Hall 1982.
[10] S.J.Farlow (editor). Self-Organizing Methods in Modeling. Marcel Deckker, 1984.
[11] G.S.Fishman. Monte Carlo Methods. Springer, 1996.
[12] P.I.Good. Resapling Methods: A Practical guide to Data Analysis. Birkhäuser, 1999.
[13] G.Infanger. Planning Under Uncertainty: Solving Large Scale Stochastic Linear Programming. Boyd and Fraser, 1994.
[14] P.Kall S.Wallace. Stochastic Programming. Wiley, 1994.
[15] D.G.Luenberger. Investment Science. Oxford, 1998.
[16] R.O.Michaud. Efficient Asset management. Harvard Business School Press, 1998.
[17] H.H.Panjer G.E.Willmot. Insurance Risk Models. Society of Actuaries, 1992.
[18] R.C.Pascual. Previsión Tecnológica de la Demanda. Marcombo Boixareu, Barcelona, 1990.
[19] S.M.Ross. Probability Models. Academic Press, 1997.
[20] J.M.Stern S.Vavasis. Active Set Algorithms for Problems in Block Angular Form. Computational and Applied Mathematics, (13) 1994.
[21] J.S.Urban-Hjorth. Computer Intensive Statistical Methods: Validadtion, Model Selec. and Bootstrap. Chapman Hall, 1994.
[22] P.Whittle. Optimization over Time. Wiley, 1982.

APPENDIX 1

```
function fus= rank2(a1,a2)

% F(t) is the component's cumulative %life
probability distribution
% F(t) = Pr(l<=t)
% Its complement is the survival probabil-
ity distribution
% Fc(t) = 1-F(t) = Pr(l>t)
% The failure probability at the next %pe-
riod x given the survival up to %current
time t is
% F(x|t) = (F(t+x)-F(t))/Fc(t)
% = 1 -Fc(x|t)
% The failure rate, hazard rate or %force
of mortality at age t is
% h(t) = f(t)/Fc(t)
% Integrating
% I[0:x] h(t)dt = -log(Fc(x))
% Fc(x) = exp(-H(x))
% H(x) = I[0:x] h(t)dt

% A(:,1)= age
% A(:,2)= h(t)

nx=100;
%maximum age at life table

% generates test assuring h(nx)==1;
% a= 1:nx; h= (1/nx)*a; h=h.^5;  %plot(h);

% f= life density; h= haz.rate;
% a= age; c=complement; u=cumulative
aux=0;
for i=1:nx
  aux= aux +h(i);
  hu(i)= aux;
  fuc(i)= exp(-hu(i));
  fu(i)= 1-fuc(i);
end

% 2 lifelong dependents
% ak= current age of k-th depend
% Xk= surviv. of k-th depend.
% R2= sup{X1,X2} R1= inf{X1,X2}
% Pr(R2<=t|a1,a2).
% = Pr(X1<=t|a1 and X2<=t|a2)
% Pr(R1<=t|a1,a2)
% = Pr(X1<=t|a1 or X2<=t|a2)
% Pr(R1>t|a1,a2)
% = Pr(X1>t|a1 and X2>t|a2)

for t=1:100
  if( (a1+t)>nx )
    fua1(t)=1;
  else      %Pr(X1<=t|a1)
    fua1(t) = ...
    ((fu(a1+t)-fu(a1))/fuc(a1));
  end
  if( (a2+t)>nx )
    fua2(t)=1;
  else
    fua2(t) = ...
    ((fu(a2+t)-fu(a2))/fuc(a2));
  end
  f2u(t)= fua1(t)*fua2(t);
  f1u(t)= fua1(t) +fua2(t) -f1u(t);
end
fus=[fu;f2u;f1u;fua1;fua2];
plot(a',fu','--b',a',f2u,'-r',a',f1u,
..'-r',a',fua1','--k',a',fua2','--k');
title( ...
..['Order statistics for survival', ...
  int2str(a1),' and ',int2str(a2)]);
```


log(hazard rate) model
-2
-4
-6
-8
40
60
80
100
hazard rate model
0.2
0.15
0.1
0.05
0
40
60
80
100
h(t) full range
1
0.8
0.6
0.4
0.2
0
40
60
80
100
survival probability
1
0.8
0.6
0.4
0.2
0
40
60
80
100



order stat. ages 25 and 50
order stat. ages 30 and 40
order stat. ages 25 and 25
order stat. ages 50 and 50
0
0.2
0.4
0.6
0.8
1
0
50
100