MARK H.A. DAVIS and MICHEL H. VELLEKOOP, London

Permanent Health Insurance: A Case Study in Piecewise-Deterministic Markov Modelling

1 Introduction

As insurance contracts become more complicated there is increasing need for a systematic approach to the associated computations, and such an approach is provided by modelling the cash flows associated with the contract as Markov processes. Of course, the idea of "markovianizing" stochastic processes by the inclusion of "supplementary variables" goes back a long way, at least to Cox [1]. This approach does not, however, seem to have been much used by actuaries for two reasons: (a) the difficulty of estimating the required transition rates and (b) the difficulty of carrying out the associated computations even when the rates are given. The former seems intrinsic: any attempt at more accurate modelling will face this problem. This paper is a contribution to overcoming (b).

There is some evidence – see for example the interesting recent paper by Jones [4] – that markovian methods are coming back into favour. There are two ways in which one can use them: one can assume a very simple process structure with a view to obtaining closed-form or easily expressed results, or one can go for a more complex (and realistic) structure and then resort to methods of numerical analysis to do the computations. We take the latter approach here. A convenient generalization of Cox's supplementary variables idea is the *piecewise-deterministic Markov processes*, introduced by the first author [2, 3], and the purpose of this paper is to examine a typical insurance application, namely permanent health insurance (PHI) [6], from this point of view in order to demonstrate that it leads to efficient and practical computational methods. Many of the ingredients are already contained in the book [3]: the Markov model is described in § 33 there, and a rather crude algorithm, based on the "general recursive method" introduced in § 32.2, is described for computing the expected payout. However, this algorithm is inefficient and there is no discussion of computation of the variance or the distribution function. In this paper we cover the latter points and introduce interpolation and extrapolation steps which give orders-of-magnitude improvement in efficiency. This does not matter much for the

simple computation of expected values but, as will be seen, the variance and – particularly – the distribution function present tougher problems and mandate a more careful examination of the solution technique. Finally, a key point in these computations is their sensitivity to parameter values; this point is not discussed at all in [3] but is covered here. Of course, this is relevant to the problem of parameter estimation in that we obtain some idea of the level of accuracy required.

PHI is a contract whereby the insured person receives payments at a fixed rate whenever he or she is sick (i.e. unavailable for work) for a time longer than a certain *elimination period.* Thus the payment made as a result of a sickness of duration d is $K \max(d-p, 0)$ where K is the contracted payout rate and p the length of the elimination period. Our objective is to compute the mean, variance and distribution of the total payout over the lifetime of the contract. The basic actuarial model is very simple and is shown in figure 1.

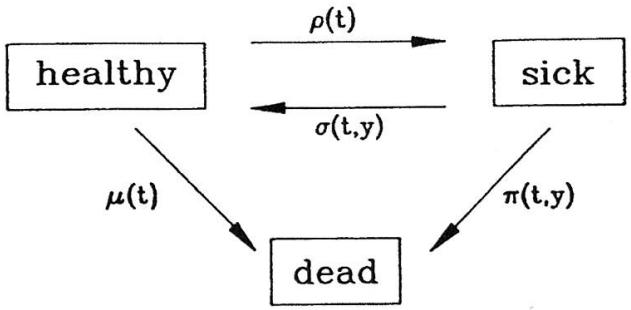

Figure 1

Transition from *healthy* to *sick* or *dead* take place with age-dependent forces $\varrho(t)$, $\mu(t)$ (here, t will denote the time since the commencement of insurance) while, when sick, the forces of transition to recovery and death are $\sigma(t,y), \pi(t,y)$ which could depend on y, the duration of the current sickness period, as well as on the age t. This paper is not concerned with the estimation problem, and we assume that the functions $\mu, \pi, \varrho, \sigma$ are known. As will be seen, models incorporating a more complicated transition structure could easily be constructed.

Section 2 below gives a brief account of the general piecewise-deterministic Markov process (PDP) and in particular describes the recursive method for solving the associated backward equation, on which all our computations are based. We return to PHI in Section 3, where the problem is formulated in terms of a PDP model.

Computational methods for calculating the mean, variance, parameter sensitivity and distribution function are given in sections 4 to 7 respectively. Section 8 contains a summary description of the algorithms together with numerical results and conclusions.

2 Piecewise-Deterministic Markov Processes

A PDP (x_t) is a random motion taking place in a state space E consisting of possibly disconnected components in some Euclidian space R^d. The process follows the integral curves of a vector field between random jumps, which may occur in the interior of E or on the boundary ∂E. The deterministic motion and jump mechanism are specified by three "local characteristics": a *vector field* $g : E \to R^d$, a *jump rate* $\lambda : E \to R_+$, corresponding to a *force of transition* in actuarial terminology, and a *transition measure* $Q : E \to \mathcal{P}(E)$, where $\mathcal{P}(E)$ denotes the set of probability measures on E. The vector field g determines a set of *integral curves* ϕ on E, i.e. $\phi(t, x)$ is the solution of the ordinary differential equation (ODE)

$$\frac{d}{dt}\phi(t,x) = g(\phi(t,x)), \qquad \phi(0,x) = x \in E \tag{2.1}$$

We can alternatively write this equation in coordinate-free form as

$$\frac{d}{dt}f(\phi(t,x)) = \mathfrak{X}f(\phi(t,x)), \qquad \phi(0,x) = x, \quad f \in C^1(E) \tag{2.2}$$

where $\mathfrak{X}$ is the first-order differential operator

$$\mathfrak{X}f(x) = \sum_{i=1}^{d} g_i(x)\frac{\partial f}{\partial x_i}(x)$$

acting on arbitrary functions $f \in C^1(E)$. (2.1) and (2.2) are equivalent ways of writing the differential equation, and the vector field operator $\mathfrak{X}$ is needed below.

Define $\Gamma = \{z \in \partial E : z = \phi(t,x)$ for some $t > 0$, $x \in E\}$ and $t_*(x) = \inf\{t : \phi(t,x) \in \Gamma\}$. Thus $t_*(x)$ is the first time the integral curve $\phi(t,x)$ hits the boundary (or $+\infty$ if it never does) and Γ is the set of boundary

points z that are hit by ϕ starting at some x in the interior of E. We call Γ the *active boundary* of E.

The PDP (x_t) follows the integral curves ϕ between random jumps, which occur at random times $0 < T_1 < T_2 \ldots$. Thus starting at $x_0 = x \in E$, $x_t = \phi(t, x)$ for $t < T_1$. The jump time T_1 is a random variable whose distribution function is

$$F(t,x) = P_x(T_1 \le t) = 1 - I_{(t<t_*(x))} \exp\left(- \int_0^t \lambda(\phi(s,x))\, ds \right),$$

implying that force of transition is $\lambda(x_s)$ on the interval $[0, t_*(x))$, with a mandatory jump at $t_*(x)$ if no jump has occurred before then. The trajectory is left-continuous at T_1, so $x_{T_1-} = \phi(T_1, x)$. The transition measure Q gives the distribution of x_{T_1} given that a jump occurs at T_1 from a position $\phi(T_1, x)$:

$$P_x[x_{T_1} \in A | T_1] = Q(A; \phi(T_1, x)) .$$

The process (x_t) jumps at time T_1 to a position selected from this distribution, and the process then restarts from x_{T_1} according to the same recipe: thus $x_t = \phi(t - T_1, x_{T_1})$ for $t \in [T_1, T_2)$, etc. It is assumed that $\lim_{n\to\infty} T_n = \infty$ a.s., a property that is easily checked in most applications such as the one in this paper.

The main general result of PDP theory is that the process just described is a homogeneous strong Markov process whose differential generator is the operator $\mathfrak{A}$, acting on functions $f : E \to R$, defined by

$$\mathfrak{A}f(x) = \mathfrak{X}f(x) + \lambda(x) \int_E (f(y) - f(x)) Q(dy; x) . \tag{2.3}$$

Sufficient conditions under which $f \in \mathcal{D}(\mathfrak{A})$ (the domain of the operator $\mathfrak{A}$) are that

$$\text{the function } t \to f(\phi(t,x)) \text{ is continuously differentiable,} \tag{2.4}$$

$$\mathbb{E}_x \sum_i |f(x_{T_i \wedge t}) - f(x_{T_i \wedge t-})| < \infty, \qquad t > 0, \quad x \in E \tag{2.5}$$

$$f(x) = \int_E f(y) Q(dy; x), \qquad x \in \Gamma . \tag{2.6}$$

Conditions (2.4) and (2.5) are technical; the main condition is the *boundary condition* (2.6). Elements z of Γ are not in the state space E and the value of $f(z)$ for $z \in \Gamma$ is *defined* by

$$f(z) = \lim_{t \downarrow 0} f(\phi(-t, z)) .$$

Thus $f(z)$ for $z \in \Gamma$ is the limit of $f(x)$ as $x \in E$ approaches z along the unique integral curve through z. The function f is in the domain $\mathcal{D}(\mathfrak{A})$ only if this limit exists and (2.6) is satisfied. The main significance of the generator is that under conditions (2.4)–(2.6) the *Dynkin formula*

$$\mathbb{E}_x[f(x_t)] = f(x) + \mathbb{E}_x \int_0^t \mathfrak{A} f(x_s)\, ds \tag{2.7}$$

holds. This is the key result from which the equations described below satisfied by expectations of various functionals are easily derived.
Let us consider the functional

$$V(x) = \mathbb{E}_x \left\{ \int_0^\infty \ell(x_t) dt + \sum_i \Phi(x_{T_i -}) I_{(x_{T_i -} \in \Gamma)} \right\}, \tag{2.8}$$

where $\ell : E \to R_+$ and $\Phi : \Gamma \to R_+$ are bounded non-negative functions. Theorem (32.10) of [3] states that if V is finite and

$$\mathbb{E}_x[V(x_t)] \to 0, \qquad t \to \infty \tag{2.9}$$

then $V(x)$ is the unique solution of the equation

$$\mathfrak{A} V(x) + \ell(x) = 0 , \qquad x \in E \tag{2.10}$$

together with the boundary condition

$$V(x) = QV(x) + \Phi(x), \qquad x \in \Gamma , \tag{2.11}$$

where $QV(x) := \int_E V(y) Q(dy; x)$. Equation (2.10) is an *integro-differential equation* involving the non-local term QV. In the application of this paper, and many others, condition (2.9) holds because of the existence of a

cemetary state; this is an isolated point $\Delta \in E$ such that $g(\Delta) = \lambda(\Delta) = 0$, so that $x_t = \Delta$ for all $t \geq \tau$, where τ is the *killing time* $\tau = \inf\{t : x_t = \Delta\}$. In our application τ is bounded by a constant and $\ell(\Delta) = 0$, so (2.9) is certainly satisfied.

Equations (2.10), (2.11) can be solved by the following iterative method ([3], § 32.2). Let ψ be an arbitrary bounded measurable function and define

$$G\psi(x) = \mathbb{E}_x\left\{ \int_0^{T_1} \ell(x_t)\, dt + I_{(T_1 = t_*(x))}\Phi(x_{T_1-}) + \psi(x_{T_1}) \right\}. \tag{2.12}$$

Iterates $G^n\psi$ are defined by $G^n\psi(x) = G(G^{n-1}\psi)(x)$, $n = 2, 3, \ldots$ and it is easily shown that

$$G^n\psi(x) = \mathbb{E}_x\left\{ \int_0^{T_n} \ell(x_t)\, dt + \sum_{i=1}^{n} I_{(x_{T_i-} \in \Gamma)}\Phi(x_{T_i-}) + \psi(x_{T_n}) \right\}$$

and hence that

$$V(x) = \lim_{n\to\infty} G^n\psi(x) \tag{2.13}$$

for any bounded measurable function ψ, where V is defined by (2.8), as long as $\mathbb{E}_x\psi(x_{T_n}) \to 0$. The function $v := G\psi$ satisfies the following modified version of (2.10), (2.11):

$$\mathfrak{A}v(x) + \lambda(x)Q\psi(x) - \lambda(x)v(x) + \ell(x) = 0\,, \qquad x \in E \tag{2.14}$$

$$v(x) = Q\psi(x) + \Phi(x) \qquad x \in \Gamma\,. \tag{2.15}$$

The point about this is that (2.14) contains no non-local terms in the unknown variable v and is in fact a first-order partial differential equation (PDE). We can thus compute V from (2.13) in which each iteration step involves only the solution of a PDE, not an integro-differential equation.

Two extensions of the above results will be needed. Firstly, we will need to compute expectations of the form

$$U(x) = \mathbb{E}_x\left\{ \sum_{i=1}^{\infty} b(x_{T_i}, x_{T_i-}) \right\}, \tag{2.16}$$

where b is some non-negative function. It is shown in [3, (31.16)] that $U(x) = V(x)$ where V is defined as in (2.8) with

$$\ell(x) = \lambda(x) \int_E b(y,x) Q(dy,x), \qquad x \in E \tag{2.17}$$

and

$$\Phi(x) = \int_E b(y,x) Q(dy,x), \qquad x \in \Gamma\,. \tag{2.18}$$

Secondly, we may want to compute net present values, i.e. discounted expectations of the form

$$V^\eta(x) = \mathbb{E}_x \left\{ \int_0^\infty e^{-\eta t} \ell(x_t)\, dt + \sum_i e^{-\eta T_i} \Phi(x_{T_i -}) I_{(x_{T_i -} \in \Gamma)} \right\}, \tag{2.19}$$

for $\eta > 0$. As discussed in [3, (31.6)] this is equivalent to "killing" the x_t process at rate η. All calculations are exactly as above, except that the generator $\mathfrak{A}$ of the process is replaced by $\mathfrak{A}^\eta f := \mathfrak{A} f - \eta f$.

3. A PDP model for PHI

The transitions in the PHI model are those shown in figure 1, but we need to record whether the duration of the current sickness period has exceeded the elimination period p or not, since payout only takes place when it has. The state space will therefore have four components $E_\nu, \nu = 1,2,3,4$, indexed by the integer variable ν:

$\nu = 1$: healthy

$\nu = 2$: sick; duration of current sickness less than p

$\nu = 3$: sick, duration of current sickness greater than p

$\nu = 4$: dead.

This gives the state space shown in figure 2, which also shows a possible trajectory of the process. E_4 is the single-point set $\{\Delta\}$.

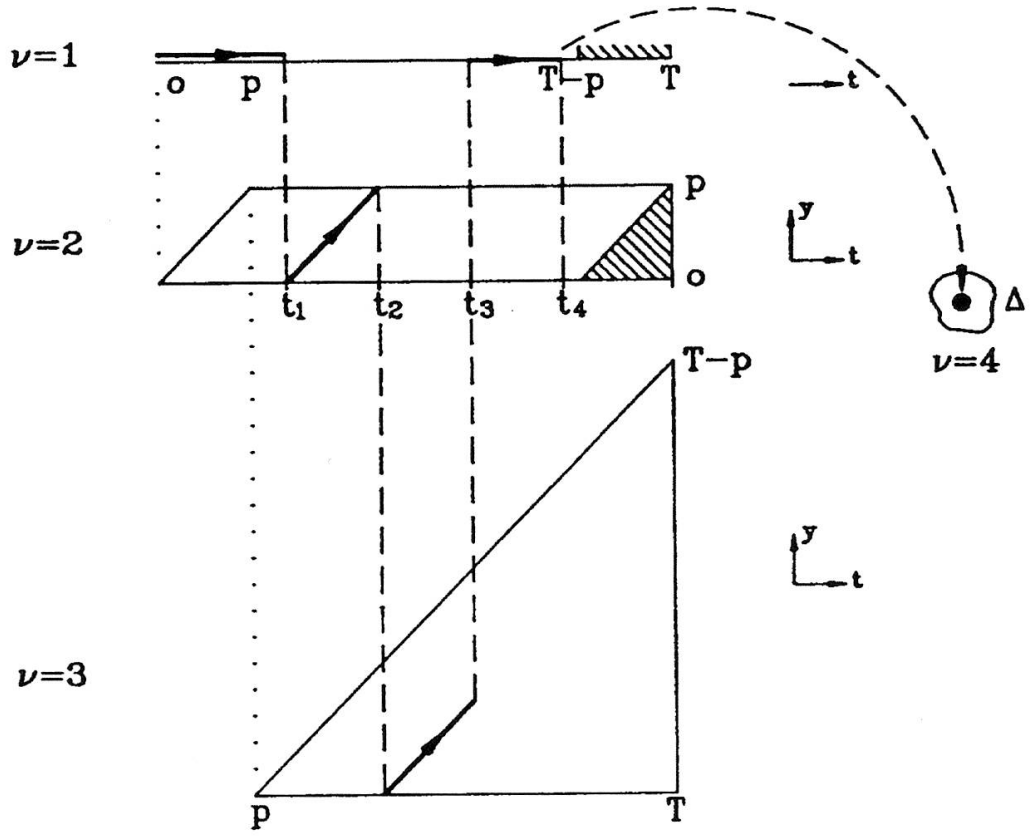

Figure 2

The horizontal axis is time t since the inception of the policy (whose duration is T). In E_2 the vertical axis variable $y, y \in [0, p]$, is the duration of the current sickness period, while the vertical variable, also denoted y, in E_3 is the excess duration of the current sickness period over the elimination period. In the sample trajectory shown, the insured person starts off healthy at time 0, and falls sick at time t_1. At time $t_2 = t_1 + p$ the sickness period reaches the elimination period and the trajectory therefore jumps into E_3, where t and $y =$ ((duration of sickness) $-p$) continue to increase at unit rate. At time t_3 the insured person recovers, but unfortunately dies at time t_4. For this trajectory the insurance company's payout would be $(t_3 - t_2)$ times the contracted payment rate (ignoring discounting; NPV calculations can also be included, as described below). A moment's thought shows that the 2-dimensional state space components E_2, E_3 contain just those (t, y) pairs that can be reached by trajectories starting in the healthy state at $t = 0$, and that the policy is valueless at all points in the shaded regions of E_1 and E_2 since, once these regions are entered, no further claims could

possibly be made. It is convenient to eliminate the shaded regions from the state space, and this can be done simply by regarding the policy as terminating at time $T-p$ in E_1; then the shaded region in E_2 will never be entered. Henceforth, E_1 and E_2 will denote these "reduced" state space components.

The state of the process will be denoted by $x=(\nu,\zeta)$ where $\zeta=t$ when $\nu=1$, $\zeta=(t,y)$ when $\nu=2$ or 3 and $\zeta=\Delta$, the cemetary state, when $\nu=4$. When $\nu=1$ the jump rate of the process is $\lambda(1,t)=\varrho(t)+\mu(t)$ and a jump from $(1,t)$ goes to $(2,(t,0))$ with probability $\varrho(t)/(\varrho(t)+\mu(t))$ and to $(4,\Delta)$ with probability $\mu(t)/(\varrho(t)+\mu(t))$, so that in the notation of the previous section the transition measure is

$$Q(dx';(1,t))=\frac{\varrho(t)}{\varrho(t)+\mu(t)}\delta_{(2,(t,0))}(dx')+\frac{\mu(t)}{\varrho(t)+\mu(t)}\delta_{(4,\Delta)}(dx')$$

where δ_x denotes the Dirac measure (unit mass) at x. We can similarly define Q in the other state space components. Writing $f(x)=f_\nu(\zeta)$, $x=(\nu,\zeta)$ for a function $f: E\to R$, and assuming as is conventional that $f(4,\Delta)=0$ we then obtain from (2.3) the following expression for the generator of the process:

$$\mathfrak{A}f_1(t)=\frac{d}{dt}f_1(t)+\varrho(t)f_2(t,0)-(\varrho(t)+\mu(t))f_1(t) \tag{3.1}$$

$$\begin{aligned}\mathfrak{A}f_2(t,y)=&\frac{\partial}{\partial t}f_2(t,y)+\frac{\partial}{\partial y}f_2(t,y)+\sigma(t,y)f_1(t)\\&-(\sigma(t,y)+\pi(t,y))f_2(t,y)\end{aligned} \tag{3.2}$$

$$\begin{aligned}\mathfrak{A}f_3(t,y)=&\frac{\partial}{\partial t}f_3(t,y)+\frac{\partial}{\partial y}f_3(t,y)+\sigma(t,y+p)f_1(t)\\&-(\sigma(t,y+p)+\pi(t,y+p))f_3(t,y)\end{aligned} \tag{3.3}$$

As regards boundary conditions, it is easiest to think of the process as being killed at time $T-p$ in E_1 and at time T in E_3 should it not have been killed earlier, i.e. transferring from $(1,T-p)$ or $(3,(T,y))$ immediately to $(4,\Delta)$ and staying there thereafter. The killing time τ is then given by $\tau=\inf\{t: x_t=\Delta\}$, as in section 2 above. (In this application, "killing" and the "cemetary state" have a gruesomely literal interpretation.) The active boundary Γ consists of the sets $\{(1,T-p)\}$, $\{(2,t,p), t\in[p,T]\}$ and

$\{(3,T,y) : y \in [0,T-p]\}$ and the PDP boundary conditions are

$$f_1(T-p) = 0\,, \tag{3.4}$$

$$f_2(t,p) = f_3(t,0), \qquad t \in [p,T]\,, \tag{3.5}$$

$$f_3(T,y) = 0, \qquad y \in [0,T-p]\,. \tag{3.6}$$

There is a small notational problem in this application: in the general theory of section 2 the process begins in some arbitrary state at time 0 and x_t denotes its state t time units later. Here, however, "real time" t is also a state space component and it is more convenient to denote by x_t the state of the process at *real time* t and to let τ be the real killing time. Thus we write

$$V(x) = \mathbb{E}_x \int_t^\infty \ell(x_s)\,ds = \mathbb{E}_x \int_t^\tau \ell(x_s)\,ds = \mathbb{E}_x \int_t^T \ell(x_s)\,ds\,, \tag{3.7}$$

with $x = (\nu,\zeta)$ and $\zeta = t$ if $\nu = 1$, $\zeta = (t,y)$ if $\nu = 2,3$, interchangeably for the expectation of an integral functional evaluated over the residual lifetime of the policy starting in state x, bearing in mind as usual that $\ell(\Delta) = 0$. We may also be interested in the NPV with discount rate η,

$$V^\eta(x) = \mathbb{E}_x \int_t^T e^{-\eta(s-t)} \ell(x_s)\,ds\,.$$

As discussed at the end of § 2, this is computed in exactly the same way as V given by (3.7), except that the generator $\mathfrak{A}$ is replaced by $\mathfrak{A}-\eta$. Referring to (3.1)–(3.3) we see that this is the same thing as replacing the forces of mortality μ,π by $\mu+\eta$, $\pi+\eta$ respectively. This occasions no change at all in the computational method, so we shall, just for notational convenience, take $\eta = 0$ throughout the rest of the paper.

4 Computation of the mean payout

Payout takes place at rate 1 per unit time per unit assured, so our objective is to compute $V(x) = V_\nu(\zeta)$, $x = (\nu,\zeta)$ given by (3.7) with

$$\ell_\nu(\zeta) = \begin{cases} 1, & \nu = 3 \\ 0 & \text{otherwise} \end{cases}$$

Then in particular $V_1(0)$ is the total expected payout over the lifetime of the policy when the insured person is healthy at the commencement of the policy. We know from the discussion in section 2 that $V(x) = \lim_{n\to\infty} G^n\psi(x)$, where $v(x) := G\psi(x)$ satisfies (2.14), (2.15) with $\Phi = 0$. Using the expressions (3.1)–(3.6) for the generator and boundary conditions, equations (2.13), (2.14) become

$$\begin{cases} \frac{d}{dt}v_1(t) + \varrho\psi_2(t,0) - (\varrho+\mu)v_1(t) = 0 \\ v_1(T-p) = 0 \end{cases} \tag{4.1}$$

$$\begin{cases} \frac{\partial}{\partial t}v_2(t,y) + \frac{\partial}{\partial y}v_2(t,y) + \sigma\psi_1(t) - (\sigma+\pi)v_2(t,y) = 0 \\ v_2(t,p) = \psi_3(t,0) \qquad t \in [p,T] \end{cases} \tag{4.2}$$

$$\begin{cases} \frac{\partial}{\partial t}v_3(t,y) + \frac{\partial}{\partial y}v_3(t,y) + \sigma\psi_1(t) - (\sigma+\pi)v_3(t,y) + 1 = 0 \\ v_3(T,y) = 0\,, \qquad y \in [0,T-p] \end{cases} \tag{4.3}$$

In these equations we have for brevity suppressed the arguments of $\varrho = \varrho(t)$ etc; these are as in (3.1)–(3.3). The first-order PDEs (4.2) and (4.3) can be solved by the "method of characteristics". The characteristic curves are the deterministic trajectories of the PDP, i.e., the lines $y \mapsto (s+y, y)$ parametrised by $s \in [0, T-p]$ in E_2 and by $s \in [p,T]$ in E_3. Along these lines v_2, v_3 satisfy the ordinary differential equations

$$\begin{cases} \frac{d}{dy}v_2^s(y) + \sigma(s+y,y)\psi_1(s+y) \\ \qquad - (\sigma(s+y,y) + \pi(s+y,y))v_2^s(y) = 0\,, \quad y \in [0,p] \\ v_2^s(p) = \psi_3(s+p,0) \end{cases} \tag{4.2'}$$

$$\begin{cases} \frac{d}{dy}v_3^s(y) + \sigma(s+y,p+y)\psi_1(s+y) \\ \qquad - (\sigma(s+y,p+y) + \pi(s+y,p+y))v_3^s(y) + 1 = 0\,, \\ v_3^s(T-s) = 0 \qquad\qquad\qquad\qquad\qquad\qquad y \in [0,T-s] \end{cases} \tag{4.3'}$$

Here $v_\nu^s(y) = v_\nu(s+y,y)$, $\nu = 2,3$. Figure 3 shows the state space with the characteristic lines.

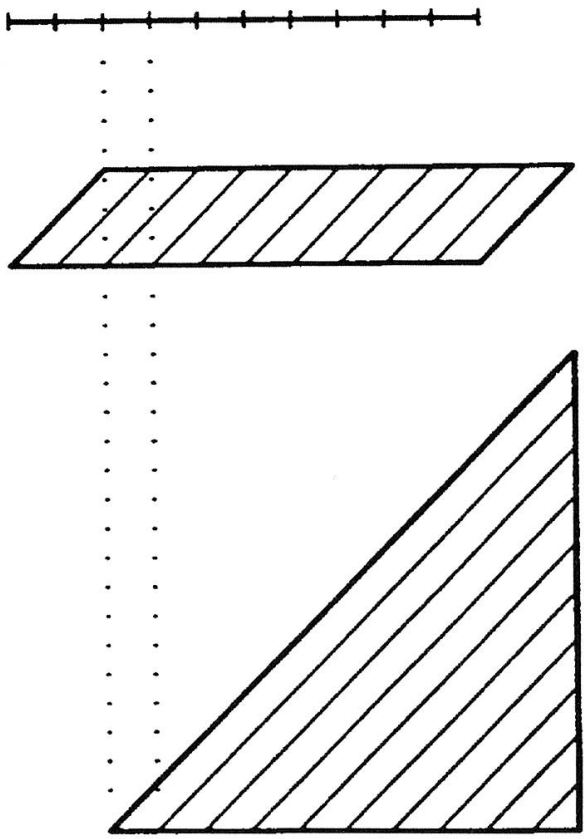

Figure 3

The value of $v(x)$ at any point x can be computed by numerical solution of (4.1), (4.2′) or (4.3′) along the characteristic line through x, in the direction of decreasing t or y with the specified boundary condition at $t = T - p$ in (4.1) or at $y = p, T - s$ in (4.2′) or (4.3′) respectively. Recall, however, that we want to compute $G\psi, G^2\psi, \ldots$, i.e. having computed $v = G\psi$ we want to re-solve (4.1), (4.2′), (4.3′) with the computed values of v replacing ψ. We therefore have to make sure that v is computed at all points needed for re-insertion into the equations at the next iteration. The easiest way to make all the computations compatible in this way is as follows:

Algorithm 1: Assume (without loss of generality) that $p = mh$ and $(T - p) = nh$ for some $h \in R$ and integers m, n. We solve (4.1) by numerical integration using the 4th-order Runge-Kutta algorithm [5] with step length h. Now (4.2′) is solved for $s = 0, h, \ldots, nh$ and (4.3′) is solved for $s = mh, (m+1)h, \ldots, (m+n)h$. To perform these integrations we are going to need the values of $\psi(x)$ at all the appropriate grid points, and at the

end of the calculation we will have obtained the values of $v(x)$ at all these points. At the next iteration we recompute, replacing the original ψ values by the computed v values. Iteration proceeds until the maximum difference between the solutions in successive iterations becomes sufficiently small.

For this small problem, Algorithm 1 executes in at most a few seconds on 486 PC or Sun workstation, but it is rather wasteful: with $n = 100$, for example, we are solving 201 differential equations per iteration. This number can be drastically reduced if one notes that the only purpose of solving the 200 equations in regions $\nu = 2,3$ is to provide the values of $\psi_2(t,0)$ needed in the next iteration of equation (4.1). Since $\psi_2(t,0)$ is a very smooth function of t we can estimate it accurately by computing a few values and obtaining the remaining values by interpolation. This leads to the following improved algorithm.

Algorithm 2: This is the same as Algorithm 1 except that (4.2′) and (4.3′) are solved only for $s = 0, qh, 2qh, \ldots, nh$ and $s = mh, (m+q)h, \ldots, (m+n)h$ respectively, where q is, say, $n/5$. We then take the computed values $v_2(0,0), v_2(qh,0), \ldots, v_2(nh,0)$ and compute a cubic spline interpolation [5] $\widehat{v}_2(t,0)$ for $t \in [0,nh]$. At the next iteration the interpolated values $\widehat{v}_2(0,0)$, $\widehat{v}_2(h,0), \ldots, \widehat{v}_2(nh,0)$ replace $\psi_2(0,0), \ldots, \psi_2(nh,0)$ in (4.1). If $n = 100$ and $q = 20$, for example, we are now solving only 11 ODEs per iteration with negligible loss of accuracy.

One could consider an alternative approach to this calculation in which the original continuous-time Markov process is approximated by a discrete-time process in which transitions are allowed only at times $0, h, 2h, \ldots$. This would lead to an algorithm not substantially different to Algorithm 1 using the Euler method for numerical integration. By staying with the continuous-time process we are able to make use of the smoothness over time of the solution to derive a substantially more efficient algorithm.

5 Computation of the variance

In previous sections we have seen how to compute $\mathbb{E}X_x$, where

$$X_x = \int_t^\infty \ell(x_s)\, ds$$

for the PDP (x_s) starting at $x = (\nu, t, y) \in E$, namely by solving (2.10), (2.11) with $\Phi(x) = 0$. We now want to consider computation of other moments of X_x, specifically the variance $\mathrm{var}(X_x) = \mathbb{E}(X_x^2) - (\mathbb{E}X_x)^2$. This is in principle easily accomplished by the stratagem of adjoining to the process an additional component (z_s), forming a new PDP $\widetilde{x}_s = (x_s, z_s)$ evolving in the state space $\widetilde{E} = \widetilde{E}_1 \cup \widetilde{E}_2 \cup \widetilde{E}_3 \cup \widetilde{E}_4$, where $\widetilde{E}_\nu = E_\nu \times R_+$, $\nu = 1, 2, 3$, and $\widetilde{E}_4 = E_4 = \{\Delta\}$. In $\widetilde{E}_1 \cup \widetilde{E}_2 \cup \widetilde{E}_3$ the component (z_s) satisfies

$$\frac{d}{ds} z_s = \ell(x_s)\,, \qquad \widetilde{x}_t = (x, z)\,,$$

so that

$$z_s = z + \int_t^s \ell(x_u)\, du\,.$$

As before, the killing time is $\tau = \inf\{t : x_t = \Delta\}$. We define $\widetilde{x}_t = \Delta$ for $t \geq \tau$, so that in particular $(\widetilde{x}_t)$ is killed at time T at the boundaries $\{T\} \times R_+$, $\{T\} \times [0, T-p] \times R_+$ of $\widetilde{E}_1, \widetilde{E}_3$ respectively, i.e. the process jumps immediately to Δ and stays there, on hitting one of these boundary sets. The union of these boundary sets is denoted $\widetilde{\Gamma}$. The new process $(\widetilde{x}_t)$ is still a PDP, and from (2.3) its generator is $\widetilde{\mathfrak{A}}$ given by

$$\widetilde{\mathfrak{A}} f(x, z) = \mathfrak{A} f(x, z) + \ell(x) \frac{\partial}{\partial z} f(x, z)\,.$$

For $\widetilde{x} = (x, z)$ define

$$b(y, \widetilde{x}) = \begin{cases} z^2, & y = \Delta, \quad \widetilde{x} \neq \Delta \\ 0, & \text{otherwise}\,. \end{cases}$$

Then clearly

$$\mathbb{E}_{\tilde{x}}\left\{\sum_{i=1}^{\infty} b(\tilde{x}_{T_i}, \tilde{x}_{T_i-})\right\} = \mathbb{E}_{\tilde{x}}(z^2_{\tau-}) = \mathbb{E}_{\tilde{x}}\left(z + \int_t^{\tau} \ell(x_s)\, ds\right)^2 .$$

We can therefore compute the latter expectation, which we denote $W(\tilde{x})$, by the method described at the end of section 2: from (2.17), (2.18) we define

$$\tilde{\ell}(\tilde{x}) = \begin{cases} z^2\mu(t), & \nu = 1\,, \\ z^2\pi(t,y), & \nu = 2\,, \\ z^2\pi(t,p+y), & \nu = 3\,, \end{cases}$$

and

$$\tilde{\Phi}(\tilde{x}) = \begin{cases} z^2, & \tilde{x} = (1, T-p, z) \quad \text{or} \quad \tilde{x} = (3, T, y, z)\,, \\ 0 & \text{elsewhere.} \end{cases}$$

Then, as described in section 2, W is the unique solution of

$$\tilde{\mathfrak{A}} W(\tilde{x}) + \tilde{l}(\tilde{x}) = 0 \qquad \tilde{x} \in \tilde{E} \tag{5.1}$$

$$W(\tilde{x}) = \tilde{\Phi}(\tilde{x}) \qquad \tilde{x} \in \tilde{\Gamma} \tag{5.2}$$

$$W_2(t,p,z) = W_3(t,0,z), \qquad (t,z) \in [p,T] \times R_+\,. \tag{5.3}$$

In particular

$$W_1(0,0) = \mathbb{E}_{(1,0)}\left(\int_0^T \ell(x_s)\, ds\right)^2 . \tag{5.5}$$

Thus the variance of the total payout starting at time 0 is just $W_1(0,0) - (V_1(0))^2$ where V_1 is as in Section 4 above. Obviously, higher moments of X_x could be computed by a similar procedure, but for concreteness we will only consider the variance here.

The first step in devising an algorithm for solving (5.1)–(5.3) is to determine the smallest subset of the new state space $\tilde{E}$ in which the solution is required. This consists of all points $(1,t,z)$, $(2,t,y,z)$ and $(3,t,y,z)$ that the trajectory $\tilde{x}_t$ could possibly reach starting at $\tilde{x}_0 = (1,0,0)$. It is

not immediately obvious what this set is, but in fact the three components are as shown in figure 4. We will in future use the notation $\widetilde{E}_\nu$, $\nu = 1,2,3$ for these reduced state space components.

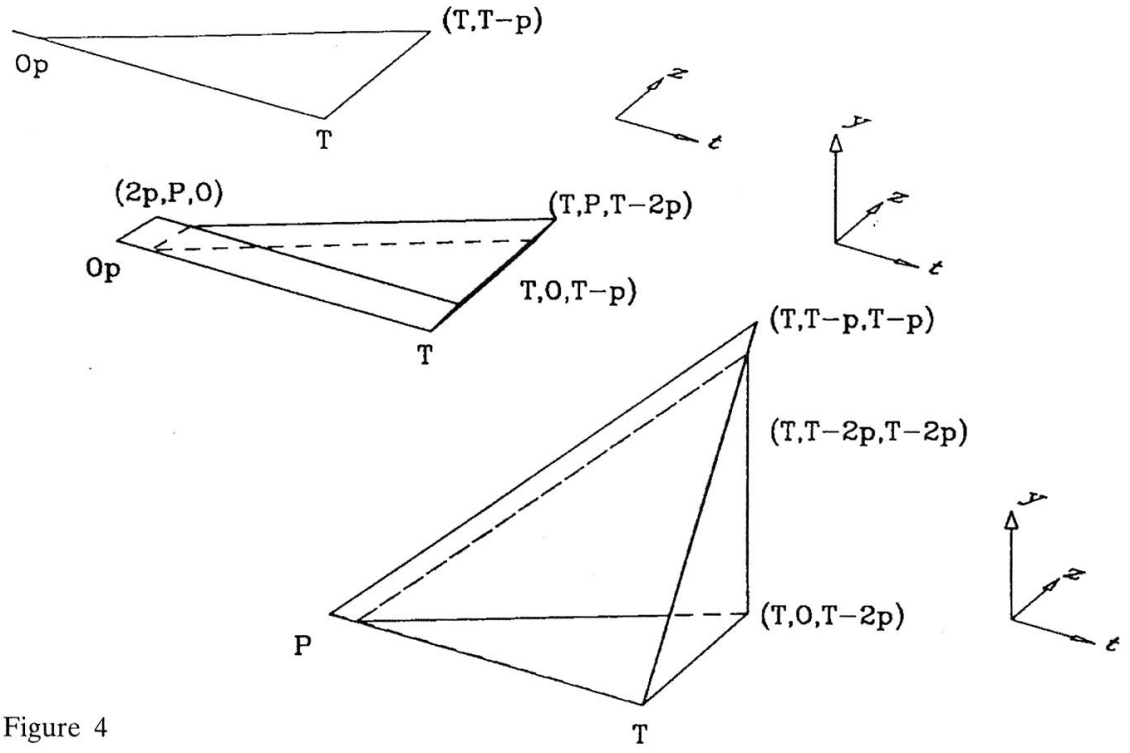

Figure 4

In principle (5.1), (5.2) can now be solved using the direct analogue of Algorithm 1, but one immediately sees that the extra dimension makes this barely feasible. In $\widetilde{E}_3$, for example, one would have to solve ODE's along characteristic lines starting at a grid of points $\{(ih, 0, jh),\ j \leq i,\ i = 0,1,2,\ldots\}$. Taking a step length $h = T/100$ there are about 5000 such points and one iteration of Algorithm 1 therefore involves numerical solution of over 10,000 ODEs! This number can however be reduced to about a dozen by the interpolation idea of Algorithm 2 and by making use of the special structure of the problem. The key point is to note that the values of $W(\widetilde{x})$ for $z > 0$ are determined by the values for $z = 0$ and the function

$$V(x) = \mathbb{E}_x \int_t^T \ell(x_s)\, ds$$

calculated in §4 above. Indeed, from (5.4) and with $\widetilde{x} = (x, z) = (\nu, \zeta, z)$, $\zeta = t$ for $\nu = 1$, $\zeta = (t, y)$ for $\nu = 2, 3$, we have

$$W(\widetilde{x}) = \mathbb{E}_{\widetilde{x}}\left(z + \int_t^T \ell(x_s)\, ds\right)^2$$

$$= z^2 + 2z\mathbb{E}_{\widetilde{x}}\left(\int_t^T \ell(x_s)\, ds\right) + \mathbb{E}_{\widetilde{x}}\left(\int_t^T \ell(x_s)\, ds\right)^2 . \tag{5.6}$$

Now observe that the two expectations in this expression do not depend on the value of z since this is simply the initial value of (z_t), which does not influence the evolution of (x_t). Hence

$$W(\widetilde{x}) = W(x, z) = z^2 + 2zV(x) + W(x, 0)\,, \tag{5.7}$$

so that the values $W(x, z)$ for $z > 0$ are determined by the values $z = 0$ together with the expectation value computed previously. Now consider the recursive computational method, which here will consist of computing

$$w(\widetilde{x}) := \widetilde{G}\psi(\widetilde{x}) = \mathbb{E}_{\widetilde{x}}\{\Phi(\widetilde{x}_{T-})I_{(T_1 \geq T)} + \psi(\widetilde{x}_{T_1})I_{(T_1 < T)}\}$$

with Φ defined as above. As before, we find that

$$w_n(\widetilde{x}) := \widetilde{G}^n\psi(\widetilde{x}) = \mathbb{E}_{\widetilde{x}}\{\psi(\widetilde{x}_{T_n})I_{(T_n < T)} + \Phi(\widetilde{x}_{T-})I_{(T_n \geq T)}\} \tag{5.8}$$

and hence, since $T_n \to \infty$ a.s., that

$$\lim_{n\to\infty} w_n(\widetilde{x}) = W(\widetilde{x})\,.$$

In detail, the equations satisfied by $w(\widetilde{x})$ are as follows. As before we write $w(\widetilde{x}) = w_\nu(\zeta, z)$ for $\widetilde{x} = (\nu, \zeta, z)$ in $\widetilde{E}$ (recall that $\zeta = t$ when $\nu = 1$, $\zeta = (t, y)$

when $\nu = 2,3$). Recalling the definitions of $\tilde{\ell}, \tilde{\Phi}$, the equations become

$$\begin{cases} \frac{\partial}{\partial t} w_1(t,z) + \varrho \psi_2(t,0,z) - (\varrho + \mu) w_1(t,z) + z^2 \mu = 0 \\ w_1(T-p,z) = z^2 \end{cases} \tag{5.9}$$

$$\begin{cases} \frac{\partial}{\partial t} w_2(t,y,z) + \frac{\partial}{\partial y} w_2(t,y,z) + \sigma \psi_1(t,z) \\ \qquad - (\sigma + \pi) w_2(t,y,z) + z^2 \pi = 0 \\ w_2(t,p,z) = \psi_3(t,0,z) \end{cases} \tag{5.10}$$

$$\begin{cases} \frac{\partial}{\partial t} w_3(t,y,z) + \frac{\partial}{\partial y} w_3(t,y,z) + \frac{\partial}{\partial z} w_3(t,y,z) \\ \qquad + \sigma \psi_1(t,z) - (\sigma + \pi) w_3(t,y,z) + z^2 \pi = 0 \\ w_3(T,y,z) = z^2 \end{cases} \tag{5.11}$$

In region 2 the characteristics are lines $y \to (s+y, y, r)$ for $y \in [0,p]$ while in region 3 they are lines $y \to (s+y, y, r+y)$, parameterized in both cases by (s,r). From (5.10), (5.11) the ODEs satisfied along characteristic lines in regions 2 and 3 are

$$\begin{cases} \frac{d}{dy} w_2^{s,r}(y) + \sigma \psi_1(s+y,r) - (\sigma + \pi) w_2^{s,r}(y) \\ \qquad + r^2 \pi(s+y,y) = 0 \\ w_2^{s,r}(p) = \psi_3(s+p,0,r) \end{cases} \tag{5.10'}$$

$$\begin{cases} \frac{d}{dy} w_3^{s,r}(y) + \sigma \psi_1(s+y,r+y) - (\sigma + \pi) w_3^{s,r}(y) \\ \qquad + (r+y)^2 \pi(s+y,p+y) = 0 \\ w_3^{s,r}(T-s) = (r+T-s)^2 \end{cases} \tag{5.11'}$$

where $w_s^{s,r}(y) = w_2(s+y,y)$ and $w_3^{s,r}(y) = w_3(s+y,y,r+y)$. If we take $\psi \equiv 0$ then from (5.8) we see that

$$w_n(\tilde{x}) := \tilde{G}^n 0(\tilde{x}) = \mathbb{E}_{\tilde{x}} \left\{ \left(z + \int_t^T \ell(x_s)\, ds \right)^2 I_{(T_n \geq \tau)} \right\}.$$

As argued before, the evolution of (x_s) (including the jump times (T_k)) does not depend on z, so that

$$w_n(\widetilde{x}) = z^2 P_x(T_n \geq \tau) + 2z\mathbb{E}_x\left\{\int_t^T \ell(x_s)\,ds I_{(T_n \geq \tau)}\right\} + \mathbb{E}_x\left\{\left(\int_t^T \ell(x_s)\,ds\right)^2 I_{(T_n \geq \tau)}\right\},$$

i.e.,

$$w_n(\widetilde{x}) = z^2 u_n(x) + 2z v_n(x) + w_n(x,0) \tag{5.12}$$

where we define $u_n(x) := P_x(T_n \geq \tau)$. The functions (u_n) can be computed in the same manner as (v_n). We define

$$u(x) = \mathbb{E}_x \sum_{i=1}^{\infty} \beta(x_{T_i}, x_{T_i -})$$

where

$$\beta(y,x) = \begin{cases} 1, & y = \Delta, \quad x \neq \Delta \\ 0 & \text{elsewhere.} \end{cases}$$

Clearly $u(x) = \lim_{n\to\infty} u_n(x) \equiv 1$ since every sample function of the process jumps to the cemetary state exactly once. Using (2.16)–(2.18) we find that $u(x) = V(x)$ given by (2.8) with $\ell_1(t) = \mu(t)$, $\ell_2(t,y) = \pi(t,y)$, $\ell_3(t,y) = \pi(t,p+y)$ and $\Phi(x) = 1$ at $x = (1, T-p)$, $(3,T,y)$, $\Phi = 0$ elsewhere. The equations for $u_n(x) := P_x(T_n \geq \tau)$ then become

$$\begin{cases} \frac{d}{dt}u_1(t) + \varrho(t)\psi_2(t,0) - (\varrho(t) + \mu(t))u_1(t) + \mu(t) = 0 \\ u_1(T-p) = 1 \end{cases} \tag{5.13}$$

$$\begin{cases} \frac{d}{dy}u_2^s(y) + \sigma(s+y,y)\psi_1(s+y) \\ \qquad - (\sigma(s+y,y) + \pi(s+y,y))u_2^s(y) + \pi(s+y,y) = 0\,, \\ \qquad\qquad y \in [0,p] \\ u_2^s(p) = \psi_3(s+p,0) \end{cases} \tag{5.14}$$

$$\begin{cases} \frac{d}{dy}u_3^s(y) + \sigma(s+y,p+y)\psi_1(s+y) \\ \qquad - (\sigma(s+y,p+y) + \pi(s+y,p+y))u_3^s(y) \\ \qquad + \pi(s+y,p+y) = 0, \qquad y \in [0,T-s] \\ u_3^s(T-s) = 1\,, \end{cases} \tag{5.15}$$

which can be solved in the same way as (v_n). Remark that $u \equiv 1$ is indeed a fixed point for these differential equations.

6 Sensitivity analysis of the mean

We know now how to approximate the expected payout $\mathbb{E}X_x$ and its variance $\mathbb{E}X_x^2$ numerically. As suggested earlier, the values obtained may depend heavily on the accuracy of the functions involved in finding them: the transition rates $\varrho(t)$, $\mu(t)$, $\pi(t,y)$ and $\sigma(t,y)$. It is therefore important to investigate the sensitivity of the expected payout to these functions. If we assume that the transition rates are continuously differentiable, it is possible to calculate the sensitivities using a numerical scheme which is very similar to the one we used for the calculation of the expected mean payout itself. We will now show how to do this for the case where these functions are constants (i.e. independent of time and the duration of the illness) but we will show later that non-constant functions can be treated in the same way. If we differentiate equations (4.1), (4.2′) and (4.3′) with respect to $\varrho(t) = \varrho$ we find, after changing the order of differentiation:

$$\begin{cases} \frac{d}{dt}\frac{\partial v_1}{\partial \varrho}(t) + \varrho\frac{\partial \psi_2}{\partial \varrho}(t,0) - (\varrho+\mu)\frac{\partial v_1}{\partial \varrho}(t) = v_1(t) - \psi_2(t,0) \\ \frac{\partial v_1}{\partial \varrho}(T-p) = 0 \end{cases} \tag{6.1}$$

$$\begin{cases} \frac{d}{dy}\frac{\partial v_2^s}{\partial \varrho}(y) + \sigma\frac{\partial \psi_1}{\partial \varrho}(s+y) - (\sigma+\pi)\frac{\partial v_2^s}{\partial \varrho}(y) = 0 \\ \frac{\partial v_2^s}{\partial \varrho}(p) = \frac{\partial \psi_3}{\partial \varrho}(s+p,0)\,, \qquad y \in [0,p] \end{cases} \tag{6.2}$$

$$\begin{cases} \frac{d}{dy}\frac{\partial v_3^s}{\partial \varrho}(y) + \sigma\frac{\partial \psi_1}{\partial \varrho}(s+y) - (\sigma+\pi)\frac{\partial v_3^s}{\partial \varrho}(y) = 0\,, \\ \frac{\partial v_3^s}{\partial \varrho}(T-s) = 0\,, \qquad y \in [0, T-s]\,. \end{cases} \tag{6.3}$$

The key observation is that the fixed point of this scheme is exactly the function

$$V_\varrho(x) = \mathbb{E}_x \int_t^\infty \ell_\varrho(x_s)\, ds$$

as in equation (3.7), where we define

$$\ell_\varrho(x) = \begin{cases} -v_1(t) + v_2(t,0) & \nu = 1 \\ 0 & \nu = 2,3 \end{cases}$$

and since ℓ_ϱ is a bounded function and $\lim_{t\to\infty} V_\varrho(x_t) = 0$, the arguments of section 2 again guarantee that this scheme converges to the unique fixed point of equations (6.1)–(6.3). This proves that we may indeed compute the sensitivities with our earlier computational method, by just replacing our old function ℓ by the new ℓ_ϱ.

It is clear that the sensitivities of the parameters $\mu(t), \sigma(t,y)$ and $\pi(t,y)$ can be found in exactly the same way, using the functions

$$\ell_\mu(x) = \begin{cases} -v_1(t) & \nu = 1 \\ 0 & \nu = 2,3 \end{cases}$$

$$\ell_\sigma(x) = \begin{cases} 0 & \nu = 1 \\ v_1(t) - v_2(t,y) & \nu = 2 \\ v_1(t) - v_3(t,y) & \nu = 3 \end{cases}$$

$$\ell_\pi(x) = \begin{cases} 0 & \nu = 1 \\ -v_2(t,y) & \nu = 2 \\ -v_3(t,y) & \nu = 3\,. \end{cases}$$

Analogously, if the parameters are not independent constants but depend on time and a parameter θ, that is

$$\varrho(t) = \varrho(t,\theta),\ \ \mu(t) = \mu(t,\theta),\ \ \pi(t,y) = \pi(t,y,\theta),\ \ \sigma(t,y) = \sigma(t,y,\theta)$$

one may simply substitute these into equations (4.1), (4.2′) and (4.3′) and differentiate with respect to θ. This will give exactly the same structure as before, when we replace the parameter functions $\varrho, \mu, \pi, \sigma$ by the derivative functions $\frac{\partial\varrho}{\partial\theta}, \frac{\partial\mu}{\partial\theta}, \frac{\partial\pi}{\partial\theta}, \frac{\partial\sigma}{\partial\theta}$, and choose an appropriate function ℓ_θ.

7 Computation of the distribution function

The method described in section 5 to compute $\mathbb{E}_x X^2$ can be modified to calculate the distribution function

$$F_x(a) = \mathbb{P}_x\left(\int_t^\tau l(x_s)\,ds \geq a\right) = \mathbb{E}_x I_{(\int_t^\tau l(x_s)\,ds \geq a)}\,.$$

In particular we can calculate $F_0(a) := F_{(1,0)}(a)$, the distribution function of the payout at time zero when healthy at the start of the policy, by computing its values for all $a \in [0, T-p]$. To do so we take, in our earlier notation

$$b(y,\widetilde{x}) = \begin{cases} I_{(z\geq a)} & y = \Delta,\ x \neq \Delta \\ 0 & \text{elsewhere} \end{cases}$$

where z is, as before, given by

$$z_s = z + \int_t^s l(x_u)\, du\,.$$

The theory in [3] gives us, as in the case of the variance

$$\widetilde{\ell}(\widetilde{x}) = \lambda(x) \int_E b(y,\widetilde{x}) Q(dy,x) = \begin{cases} I_{(z\geq a)}\mu(t) & \nu = 1 \\ I_{(z\geq a)}\pi(t,y) & \nu = 2 \\ I_{(z\geq a)}\pi(t,p+y) & \nu = 3 \end{cases}$$

and boundary conditions

$$\Phi(\widetilde{x}) = \begin{cases} I_{(z\geq a)} & \widetilde{x} = (1, T-p, z) \text{ or } \widetilde{x} = (3, T, y, z) \\ 0 & \text{elsewhere.} \end{cases}$$

Defining

$$d_n(a,\widetilde{x}) = d_n(a,x,z) = \mathbb{P}_{\widetilde{x}}\left(z + \int_t^{T_n} l(x_s)\, ds \geq a\right)$$

as the n-th estimate of $F_x(a)$ starting from a point $\widetilde{x}$ in $\widetilde{E}$, this leads us to the following equations, after transforming to characteristic coordinates

$$\begin{cases} \frac{d}{dt} d_1(a,t,z) + \varrho(t)\psi_2(a,t,0,z) - (\varrho(t)+\mu(t)) d_1(a,t,z) \\ \qquad\qquad + I_{(z\geq a)}\mu(t) = 0 \\ d_1(a, T-p, z) = I_{(z\geq a)} \end{cases} \tag{7.1}$$

$$\begin{cases} \frac{d}{dy} d_2^{s,r}(a,y) + \sigma(s+y,y)\psi_1(a,s+y,r) - \\ \qquad (\sigma(s+y,y) + \pi(s+y,y)) d_2^{s,r}(a,y) \\ \qquad + I_{(r \geq a)} \pi(s+y,y) = 0 \\ d_2^{s,r}(a,p) = \psi_3(a,s+p,0,r)\,, \qquad y \in [0,p] \end{cases} \tag{7.2}$$

$$\begin{cases} \frac{d}{dy} d_3^{s,r}(a,y) + \sigma(s+y,p+y)\psi_1(a,s+y,r+y) - \\ \qquad (\sigma(s+y,p+y) + \pi(s+y,p+y)) d_3^{s,r}(a,y) \\ \qquad + I_{(r+y \geq a)} \pi(s+y,p+y) = 0\,, \\ d_3^{s,r}(a,T-s) = I_{(r+T-s \geq a)}\,, \qquad y \in [0,T-s]\,. \end{cases} \tag{7.3}$$

As in the previous case, we only have to do the actual computations for $z = 0$ and $r = 0$ in $\widetilde{E}_1$ and $\widetilde{E}_2, \widetilde{E}_3$ respectively since

$$d_n(a,x,z) = d_n(a-z,x,0)\,. \tag{7.4}$$

The only expression which involves a term with the third coordinate not equal to zero is $\psi_1(a,s+y,r+y) = \psi_1(a-y,s+y,0)$ in (7.3). We should therefore compute the distribution function recursively for $a = 0, h, 2h, \ldots, nh$, with the same step length h as we use for the integration of equations (7.1)–(7.3), storing previously calculated values of $d_1(a,kh,0)$, $k = 0, \ldots, n$ in memory for use in later iterations. Note that this means that we have to solve equations (7.1)–(7.3) for all these values of a so the computation of the distribution function requires n times the number of computations we had to do for the mean or variance! We have not been able to find any method for computing $F_0(a)$ in larger increments of a. However, this is not very surprising since computation of the distribution function can be interpreted as calculating *all* moments of the payout, so this increase in computational effort is only natural.

8 Implementation and numerical examples

We can now give the algorithm for the computation of mean, variance and sensitivity functions:

Algorithm 3: As before we suppose that $p = mh$ and $T - p = nh$ for integers n, m where h is the step length for numerical integration, and we

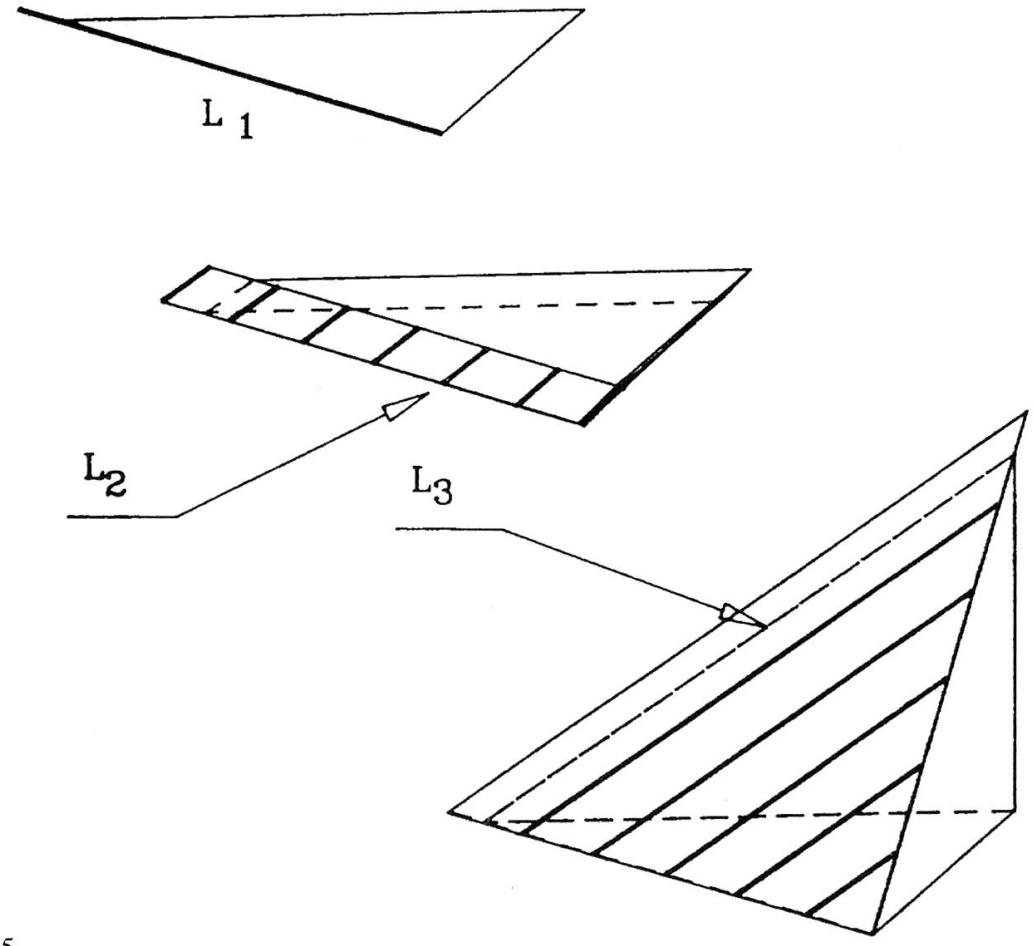

Figure 5

take further integers q, n', m' such that $q = n/n' = m/m'$ (n' being typically between 15 and 20). Suppose that $w_n(\widetilde{x}), v_n(x), u_n(x)$ have been computed in the previous iteration or, for $n = 0$, $w_0 = v_0 = u_0 \equiv 0$. To find the numerical solution of the equation for $w_n(\widetilde{x})$ we want to compute its values in the following points on the characteristic lines, as displayed in figure 5:

$$\begin{aligned}
&\widetilde{E}_1 \supset \{(1,t,0) \mid t = kh, \ \ k = 0, \ldots, n\} \\
&\widetilde{E}_2 \supset \{(2,s+y,y,0) \mid s = kqh, \ \ y = jqh, \ \ k = 0, \ldots, n', \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = 0, \ldots, m'\} \\
&\widetilde{E}_3 \supset \{(3,s+y,y,y) \mid s = kqh, \ \ y = jqh, \ \ k = m', \ldots, m'+n', \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = 0, \ldots, m'+n'-k\}
\end{aligned}$$

This leads to the following algorithm:

1. Compute $u_{n+1}(x)$ by substituting the values of $u_n(x)$ in equations (5.13)–(5.15) and using a 4th order Runga Kutta scheme [5] to solve the differential equations.

2. Compute $v_{n+1}(x)$ in the same way by substituting the values of $v_n(x)$ in equations (4.1), (4.2′) and (4.3′).

3. (Interpolation Step) Use the values of $w_{n,2}(kqh, 0, 0)$, $k = 0, \ldots, n'$ to compute a cubic spline interpolate $\widehat{w}_{n,2}(t, 0, 0)$, $0 \leq t \leq T - p$.

4. (Extrapolation Step) Calculate the values of $w_{n,1}(p + kh, jh)$, $k = 0, \ldots, n$, $j = 0, \ldots, k$ using equation (5.12).

5. Substitute the values obtained in the two previous steps together with the values of $w_n(\widetilde{x})$ in equations (5.9), (5.10′) and (5.11′) to compute $w_{n+1}(\widetilde{x})$, taking $z = 0$ in (5.9) and $r = 0$ in (5.10′) and (5.11′). Use a step length h in $\widetilde{E}_1$ and a step length qh in $\widetilde{E}_2 \cup \widetilde{E}_3$.

6. Repeat this until the maximum differences $|v_{n+1}(x) - v_n(x)|$ and $|w_{n+1}(\widetilde{x}) - w_n(\widetilde{x})|$ are both below some positive threshold ε.

7. Substitute the final values of $v_n(x)$ in the sensitivity function ℓ_ϱ defined in equation (6.5) (or in $\ell_\mu, \ell_\pi, \ell_\sigma$, depending on which sensitivity one wants to compute) and solve the equations (6.1)–(6.3) to obtain the iterates $\left(\frac{\partial v}{\partial \varrho}\right)_n(x)$, taking them all zero during the first iteration. Repeat this until the maximum difference $\left|\left(\frac{\partial v}{\partial \varrho}\right)_{n+1}(x) - \left(\frac{\partial v}{\partial \varrho}\right)_n(x)\right|$, is below the threshold ε.

The stopping criterion mentioned in 6. can be improved to guarantee a certain accuracy of the estimates. Indeed, since we start the algorithm with $\psi_0 \equiv 0$ we have for the error in the expected payout at time zero after n iterations:

$$\begin{aligned}
\left|V_1(0) - [G^n\psi_0]_1(0)\right| &= \left|\mathbb{E}_0 \int_0^\tau l(x_s)\,ds - \mathbb{E}_0 \int_0^{T_n} l(x_s)\,ds\right| \\
&= \mathbb{E}_0 \int_{T_n}^\tau l(x_s)\,ds \cdot I_{(T_n \le \tau)} \\
&\le \mathbb{E}_0 \int_0^{T-p} 1\,ds \cdot I_{(T_n \le \tau)} \\
&= (T-p)\mathbb{P}_0(T_n \le \tau) \\
&= (T-p)(1-u_n(0))
\end{aligned}$$

where we have used the definition of $u_n(x)$ to obtain the last equality. Completely analogously one may derive that

$$\left|W_1(0,0) - [G^n\psi_0]_1(0)\right| \le (K(T-p)^3 + (T-p)^2)(1-u_n(0))$$

where K denotes the maximum of $\mu(t)$, $\pi(t,y)$ for $0 \le t \le T$ and $0 \le y \le T$, and therefore an upper bound $\delta > 0$ on the errors in both the mean and variance is guaranteed if we replace step 6. by

6. Repeat this until $u_n(0) > 1 - \dfrac{\delta}{K(T-p)^3 + (T-p)^2}$.

Remark that this stopping criterion makes sure that the algorithm will indeed stop after a finite number of iterations since $u_n(0) \to 1$ when $n \to \infty$. The algorithm for the calculation of the distribution function becomes

Algorithm 4: For all values of $a = 0, h, 2h, \ldots, nh$, solve equations (7.1)–(7.3) recursively with integration step length h and with $z = 0$ in (7.1) and $r = 0$ in (7.3). Take $\psi_0 \equiv 0$ during the first iteration and use formula (7.4) to calculate the term $\psi_1(a, s+y, y)$ in (7.3). Stop iteration as soon as

the iterates change less than a certain prescribed bound ε, and then store the values of $d_1(a, kh, 0)$, $k = 0, \ldots, n$ in memory for later use in (7.4). Continue with the next value of a, until all values have been calculated. Alternatively, to guarantee a maximal error of δ for $\|F_0 - [G^n \psi_0]_1(0)\|_{\max}$, also calculate the functions $u_n(x)$ using (5.13)–(5.15) during the iterations for a particular value of a, and stop iteration when $u_n(0) > 1 - \delta$. It is easily shown by the method outlined above that this guarantees that the error will be smaller than δ.

We will now discuss some examples and results obtained by implementing the algorithms 3 and 4.

Example 1: A good test case for algorithms 3 and 4 consists of a model in which the probabilities of dying and recovering are zero while there is a constant positive probability of falling ill,

$$\varrho(t) = \varrho > 0, \qquad \mu(t) = \pi(t, y) = \sigma(t, y) = 0$$

since we can explicitly derive the solution for this case. Indeed, $T - p - X_0$ is approximately exponentially distributed with parameter $\frac{1}{\varrho}$, since we expect to fall ill after a time $\frac{1}{\varrho}$ and remain ill the rest of the time $T - \frac{1}{\varrho}$ since there is no chance of recovery. Therefore the expected payout at time 0 is $T - p - \frac{1}{\varrho}$ and both the variance and the sensitivity function of ϱ should be $\frac{1}{\varrho^2}$. The distribution function of the payout (measured in years) is approximately $u \to e^{\varrho(u-(T-p))}$.

The results are presented in figure 6 which shows the numerical approximation for the value $\varrho = 0.3$ and $T = 25.0$, $p = 0.5$ with a step length $h = \frac{T}{400}$.

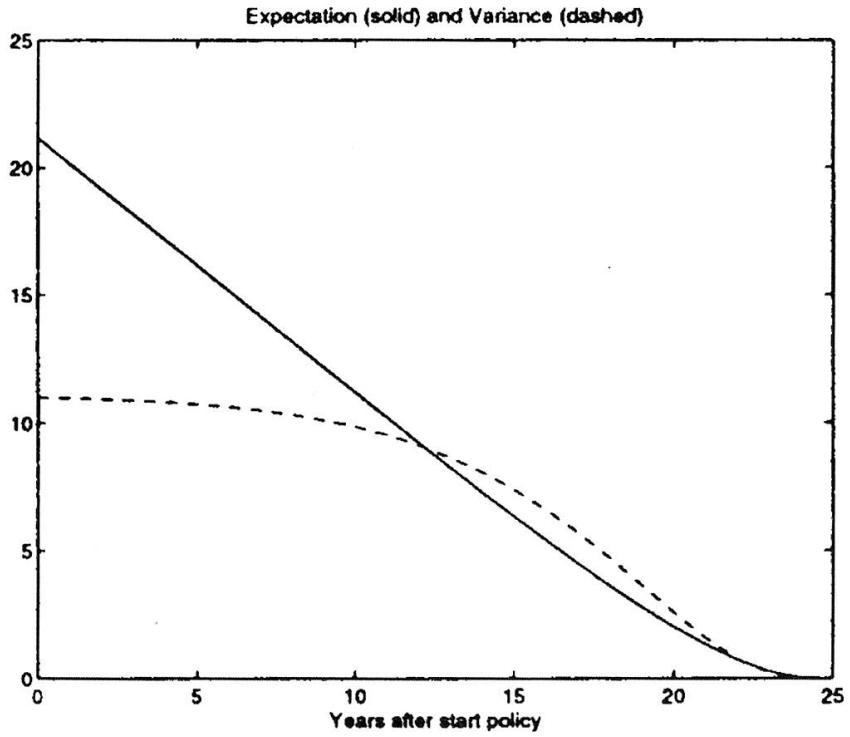

Figure 6: Example 1, Expectation and Variance

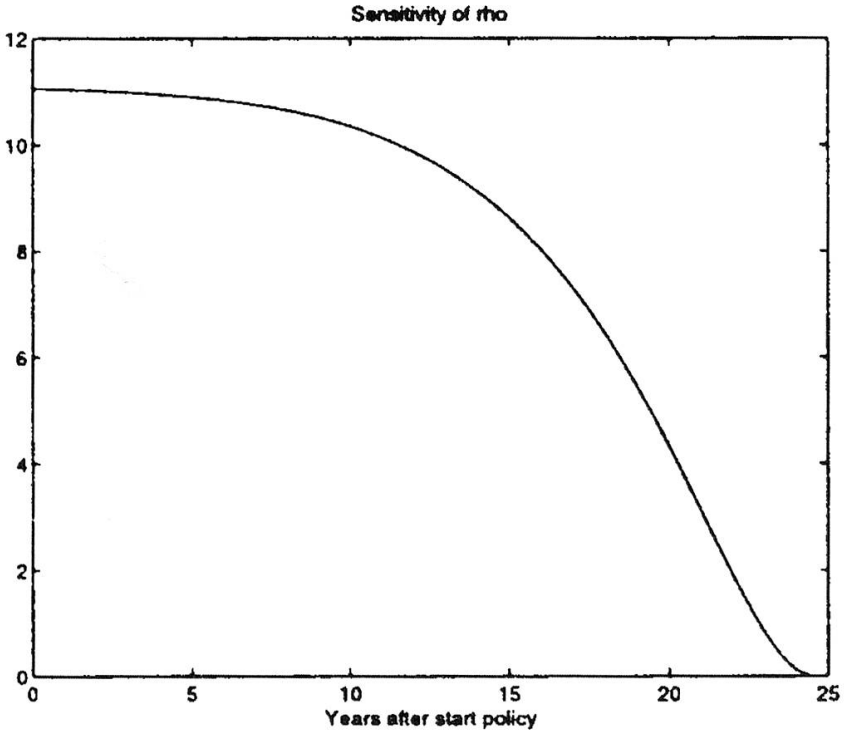

Figure 7: Example 1, Sensitivity of ϱ

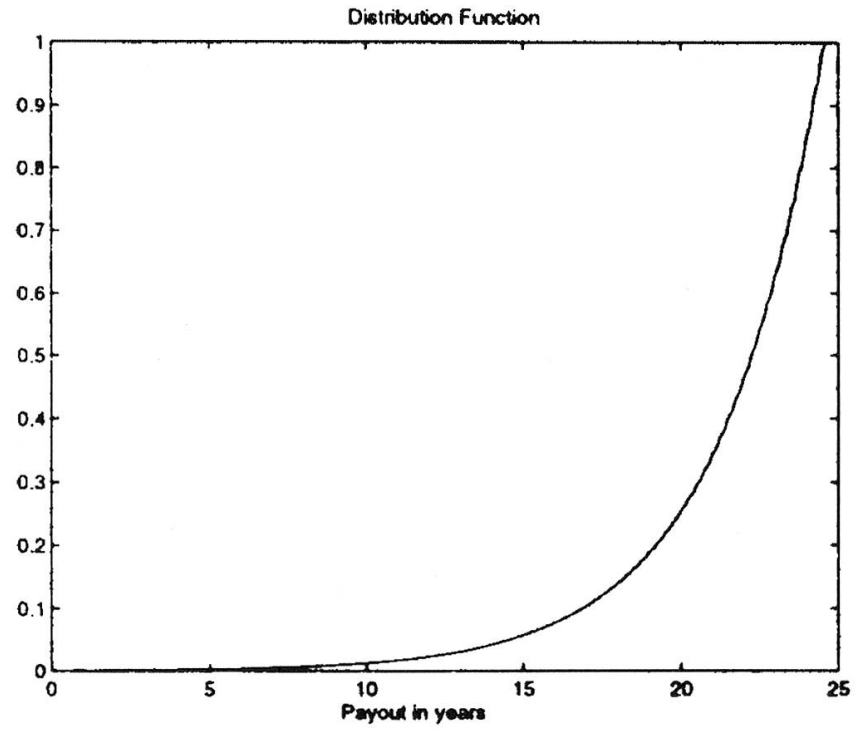

Figure 8: Example 1, Distribution.

There is excellent agreement after only four iterations for the mean, variance, and sensitivity which for this case should have the values

$$\begin{aligned}
\mathbb{E}X_0 &= T - p - \frac{1}{\varrho} = 21.2 \\
\mathrm{Var}X_0 &= \frac{1}{\varrho^2} = 11.1 \\
\frac{\partial}{\partial \varrho}\mathbb{E}X_0 &= \frac{1}{\varrho^2} = 11.1
\end{aligned}$$

and the distribution function, shown in figure 8, is as expected.

Example 2: To compare the sensitivity functions of the different parameters a model was used in which all transition rates are constant:

$$\begin{aligned}
\varrho(t) &= 0.3 \\
\mu(t) &= 0.01 \\
\pi(t,y) &= 0.01 \\
\sigma(t,y) &= 2.8
\end{aligned}$$

The results of the computations, which used the same step length, insurance period and elimination period as before, are presented in figures 9 to 13. Apart from the mean and variance in the first figure, the *relative sensitivity functions*

$$\frac{\eta}{V_1(x)}\left(\frac{\partial V_1(x)}{\partial \eta}\right), \qquad \eta = \varrho, \mu, \pi, \sigma$$

are shown. Note the difference between the sensitivity of the dying transitions and the other two. The results clearly suggest that an accurate estimation of the parameter σ is much more important than the estimation of the other parameters and that the data concerning the probabilities of dying are far less significant. Of course, the sensitivities of μ, π and σ are negative, since a larger probability of death or recovery from illness reduces the expected payout.

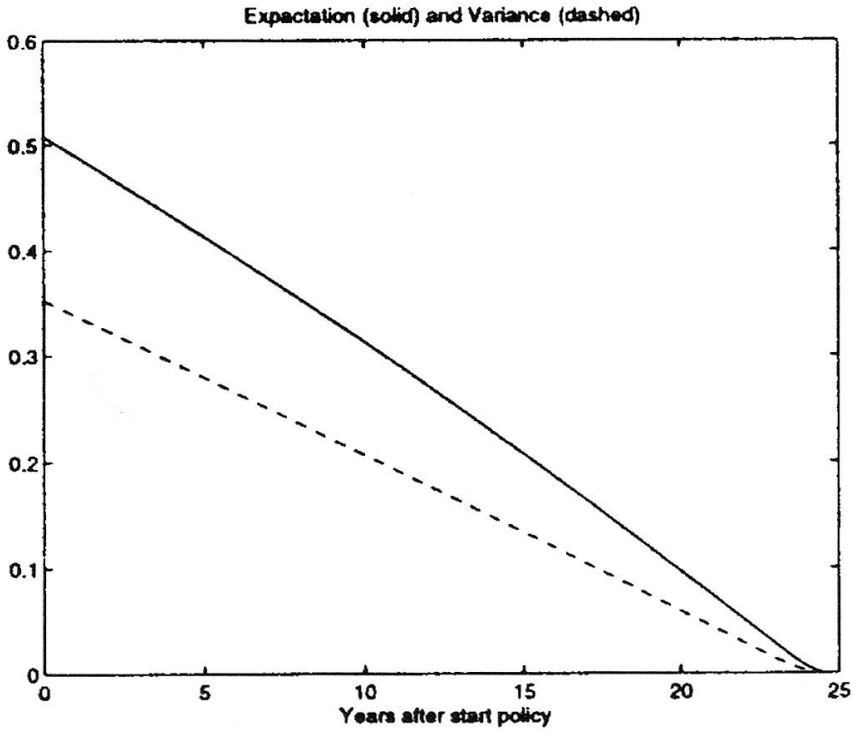

Figure 9: Example 2, Expectation and Variance

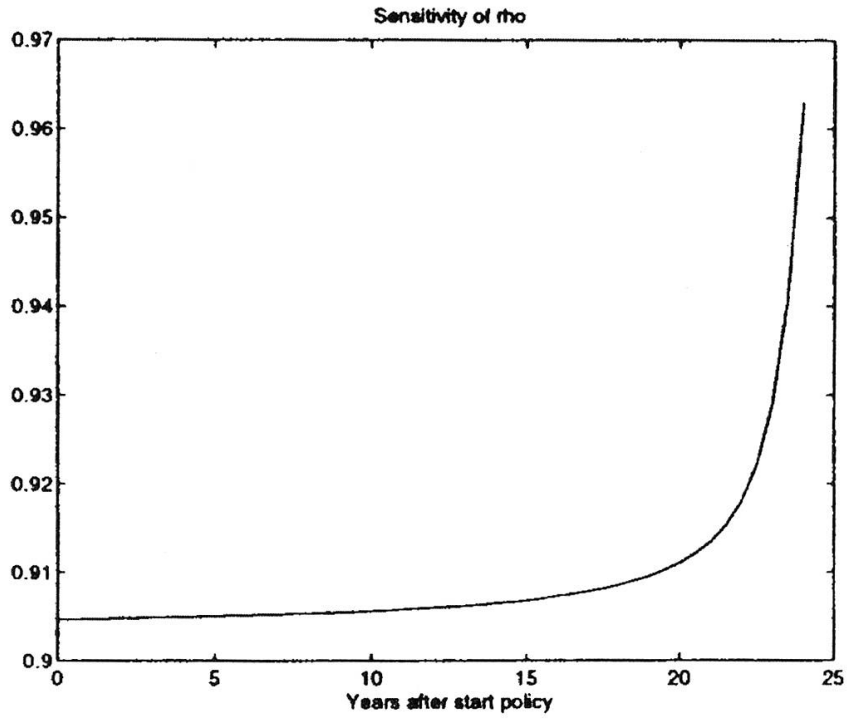

Figure 10: Example 2, Sensitivity of ϱ

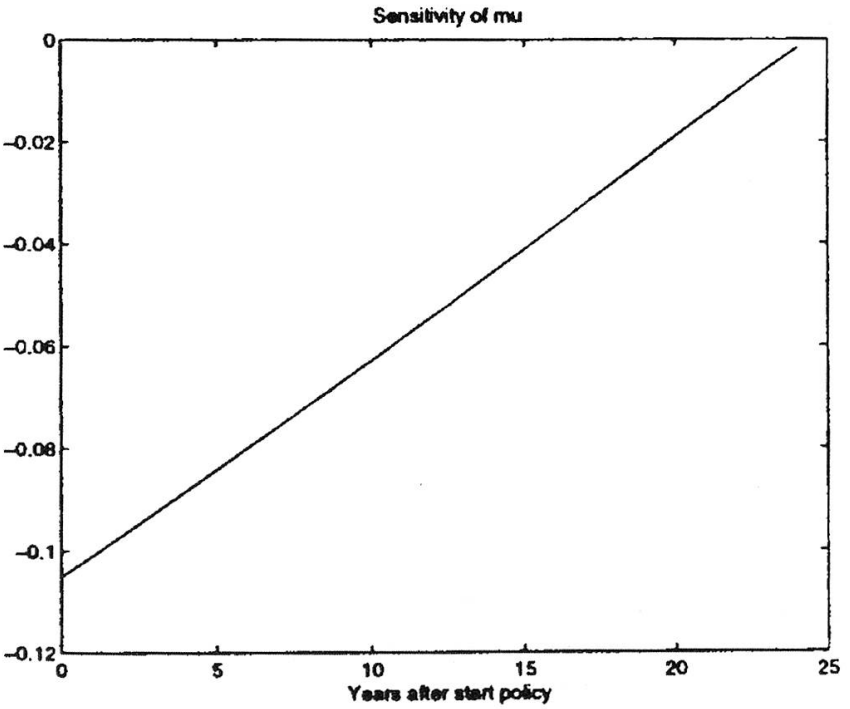

Figure 11: Example 2, Sensitivity of μ

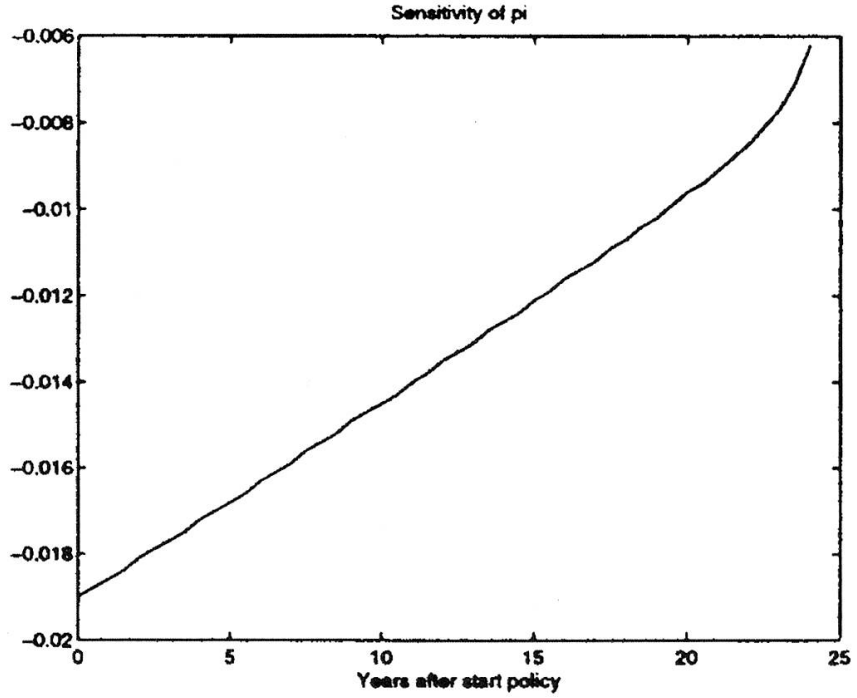

Figure 12: Example 2, Sensitivity of π

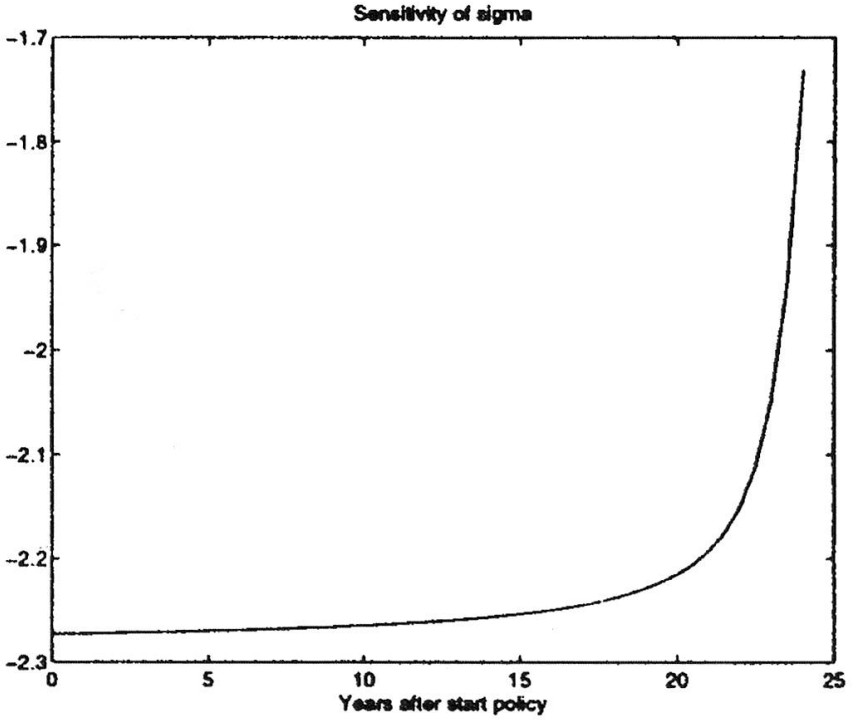

Figure 13: Example 2, Sensitivity of σ

Example 3: In a more complicated model we took

$$\begin{aligned}
\varrho(t) &= 0.3 + 0.1k(t) \\
\mu(t) &= 0.01 \\
\pi(t,y) &= 0.01(1 + 2(1 - e^{-y})) \\
\sigma(t,y) &= 2.8(1 - e^{-2y})(1 - 0.2k(t)) \\
k(t) &= \begin{cases} 0 & 0 \le t \le 15 \\ 0.01(t-15)^2 & 15 < t \le T \end{cases},
\end{aligned}$$

This means that the death rate when ill grows slowly from the normal death rate to three times the normal death rate, that the chances of falling ill rise slowly after 15 years from 0.3 to 0.4 and that the probability of recovery is mildly decreasing in the same way and rises from zero to its final value after roughly one year. The results of this simulation for $T = 25.0$ years, elimination period $p = 0.5$ years and step length $h = \frac{T}{400}$ are presented in the figures 14 and 15. The mean and variance converged after 44 iterations; some intermediate results are shown as well.

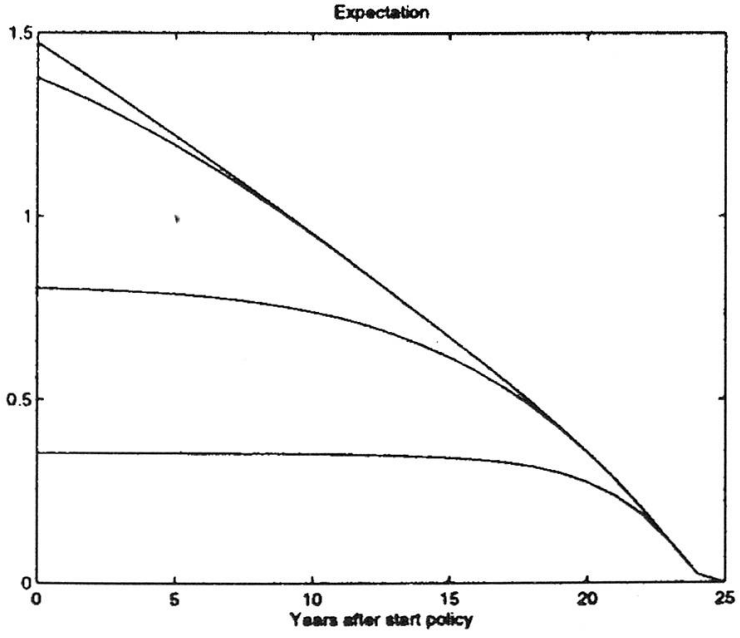

Figure 14: Example 3, Expectation

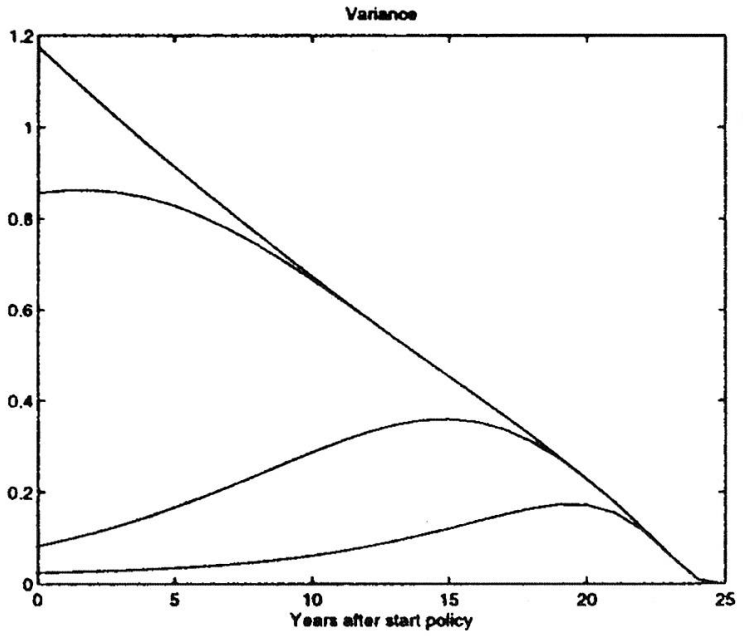

Figure 15: Example 3, Variance

In figure 16 the distribution function of the payout at time $t = 0$ is shown. The mean and variance for this distribution function can be found by numerical integration. They are approximately $\mathbb{E}X_0 \approx 1.9$ and $\mathrm{Var}X_0 \approx 1.4$, which agrees nicely with the values found for $t = 0$ in the calculations of the mean and variance.

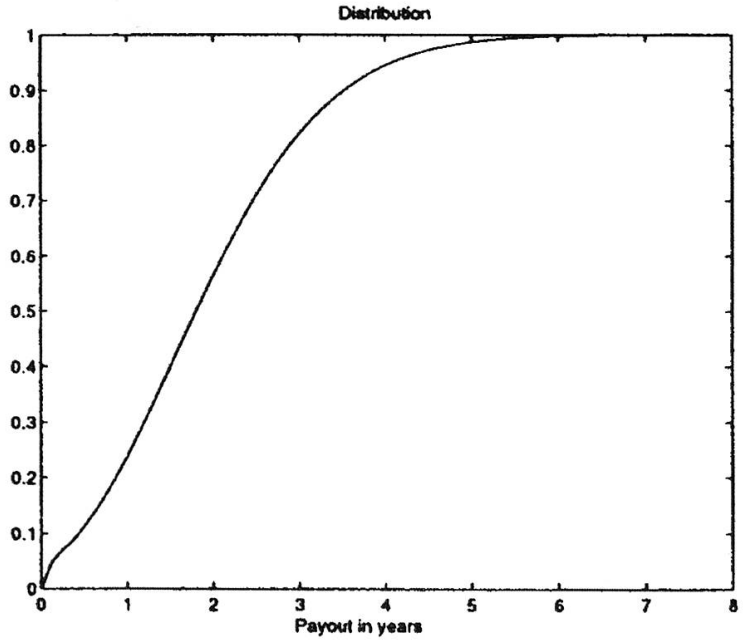

Figure 16: Example 3, Distribution

Acknowledgement. The authors are grateful to the referee for some thoughtful comments on the original version of this paper.

References

[1] *D.R. Cox*, The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables, Proc. Cam. Phil. Soc. **51** (1955) 433–441

[2] *M.H.A. Davis*, Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models (with discussion), J. Royal Statist. Soc. (B) **46** (1984) 353–388

[3] *M.H.A. Davis*, Markov Models and Optimization, Chapman & Hall, London, 1993

[4] *B.L. Jones*, Actuarial calculations using a Markov model, Transactions of the Society of Actuaries, Vol. XLVI (1995) 23–46

[5] *W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling*, Numerical Recipes in C (2nd ed.), Cambridge University Press, New York, 1992

[6] *H. Waters et al.*, Permanent Health Insurance, CMI Report 12, Institute of Actuaries, London, 1991

Mark H.A. Davis
Mitsubishi Finance International plc
6 Broadgate
London EC2M 2AA, UK

Michel H. Vellekoop
Centre for Process Systems Engineering
Imperial College
London SW7 2BY, UK

Summary

Permanent health insurance is a contract whereby the insured person is paid at a fixed rate whenever he/she is sick for longer than a minimum *elimination period*. The payout process resulting from this contract is modelled as a piecewise-deterministic Markov process, and algorithms are presented by which the mean, variance and indeed the distribution function of the payout may be computed. These algorithms involve iterative solution of systems of first-order partial differential equations, and this is accomplished by combining the "method of characteristics" with certain interpolation and extrapolation steps. Numerical results are given.

Zusammenfassung

Der Artikel betrachtet Verträge, gemäss denen die versicherte Person einen festen Betrag pro Zeiteinheit ihrer Krankheit erhält, nachdem eine minimale *Wartefrist* verstrichen ist. Der aus dem Vertrag resultierende Leistungs-Prozess wird als stückweise deterministischer Markovprozess modelliert. Dabei werden Algorithmen angegeben, mit welchen der Erwartungswert, die Varianz und – in der Tat – die Verteilungsfunktion der Leistungen bestimmt werden können. Diese Algorithmen benutzen iteratives Lösen von Systemen partieller Differentialgleichungen erster Ordnung, und dazu wird die "Charakteristiken-Methode" mit gewissen Inter- und Extrapolationsschritten kombiniert. Numerische Resultate werden angegeben.

Résumé

Ce travail traite des contrats d'assurance selon lesquels la personne assurée touche un montant fixe par unité de durée de maladie mesurée à partir d'un *délai d'attente*. Le processus des prestations résultant de ces contrats est décrit à travers un processus de Markov déterministe par morceaux. On donne des algorithmes permettant de déterminer la moyenne, la variance et – en fait – la fonction de répartition des prestations. Ces algorithmes requièrent la résolution par itérations de systèmes d'équations différentielles partielles de premier ordre. Cette dernière est effectuée en combinant la méthode des caractéristiques avec certains pas d'interpolation et d'extrapolation. On donne des résultats numériques.