Chapter V

Multi-state insurances

Up to now we dealt with one state which was called a status–alive), from which we passed to some states, called status–death caused by a "decrement". Thus at the exit of the status, the insurance ended and the payout dependent on the state defined by the risk. In many cases, situation can be more complicated. We now give few examples, the first one deals with cases considered in former chapters. Suppose we consider a life aged of (x) and at time t his/her state of life is $\mathcal{J}_x(t) = \mathcal{J}(t)$, where $\mathcal{J}(t)$ evolves among states.

Example 0.1 Consider the case with one life and one risk. Let $T = T_x$ be the future lifetime. We have two states "ALIVE"- 1 and "DEAD"- 0. Sometimes we will abbreviate "ALIVE" by A and "DEAD" by D. Then

$$\mathcal{J}_x(t) = \begin{cases} \text{ALIVE} - -1 & \text{if } T_x \leq t, \\ \text{DEAD} - -0 & \text{if } T_x > t. \end{cases} \tag{0.1}$$

This model we call "ALIVE–DEAD", in short AD. Similarly when $K = K_x$ is curtate future lifetime of life aged (x), we define a discrete time process

$$\mathcal{J}_{[x]+k} = \begin{cases} 0 & \text{if } K_x < k, \\ 1 & \text{if } K_x \geq k\,, \end{cases} \tag{0.2}$$

$k = 1, \ldots$. We have to fix $\mathcal{J}_{[x]+0} = 1$ since at the beginning (x) is alive. If (x) dies in the first year of the policy, then $\mathcal{J}_{[x]+1} = 0$. [1] Consider now one life and few decrements. That is we have (T_x, J_x) as it was in Chapter

[1] W tym przykadzie $\mathcal{E}_{\text{abs}} = \{0\}$, $\mathcal{E}_{\text{tra}} = \{1\}$. Na przykad ubezpieczenia/renty rozpatrywane w rozdziaach od 2 do 3 s zwizane z przejciem procesu $\mathcal{J}(t)$ ze stanu 1 do stanu 0 lub

IV defined by mortality rates $\mu^{(j)}(t)$ $(j = 1, \ldots, k)$. Then we introduce the process $\mathcal{J}_x(t)$ przyjmujcy wartoci $0_1, ..., 0_k$,1 gdzie

$$\mathcal{J}_x(t) = \begin{cases} \text{ALIVE} - 1 & \text{if } T_x > t \\ \text{D}_1 - 0_1 & \text{if } T_x \le t, J_x = 1 \\ \text{D}_2 - 0_2 & \text{if } T_x \le t, J_x = 2 \\ \cdots & \cdots \\ \text{D}_k - 0_k & \text{if } T_x \le t, J_x = m \end{cases} \tag{0.3}$$

[2]

We can make the following classification of states: *absorbing* ones (thatis if the process enters an absorbing state, then it stays in this state for ever) and the set of such the states we denote by $\mathcal{E}_{\text{abs}}$, and *transient* denoted by $\mathcal{E}_{\text{tra}}$. In examples above states D and $D_1, \ldots, D_k$ are absorbing and A is transient. For mathematical convenience we number states by $0_1, \ldots, 0_k, 1, \ldots, l$. Unless otherwise stated, state 0 is the only absorbing one. The space of states we denote by $\mathcal{E} = \mathcal{E}_{\text{abs}} \cup \mathcal{E}_{\text{tra}}$.

In multistate insurances we can have payout dependent on being in a state in a form of an annuity, and by transition between states in form of lump payment. As usual time t runs over $[0, n]$. We notice that $\mathcal{J}(t)$ is a stochastic process, correspondingly $\mathcal{J}_k$ a sequence of random variables. From now on we are supposing that realizations of $\mathcal{J}(t)$ are piecewise constant with finite number of jumps in $[0, n]$ and at jump moment right continuous with lefthand limits.

Example 0.2 We now consider models with three states. We have two variants: “HEALTHY”, “DISABLED” and “DEAD” in the *permanent disability model* or “HEALTHY”, “SICK” and “DEAD” in the *disability income insurance model*. In the first case a policy provides some of the following benefits: an annuity while disabled, a lump sum on becoming permanently disabled, a lump sum on death, and premium paid in a form of an annuity while healthy.

---

od stanu 1 do $0_1, \ldots, 0_k$. Natomiast co si dzieje dalej w wczeniejszych rozdziaach ju nas nie interesowao. Ten model bdziemy nazywa *ycie – mier*. Jednake istniej polisy, z ktrych wypaca si wiadczenia po mierci ubezpieczonego; patrz zadanie 7.3 gdzie rozpatruje si tzw. ubezpieczenie zaopatrzenia rodzinnego [family income insurance].

[2] $\mathcal{E}_{\text{abs}} = \{0_1, ...0_k\}$.

Transitions in such the model are depicted in Figure 0.2. We leave to the reader to describe possible benefits for the disability income insurance model.[3]

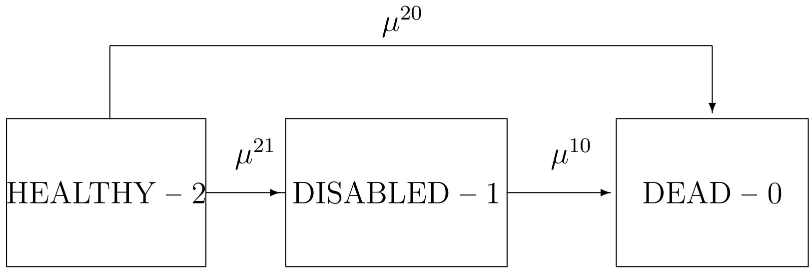

Figure 0.1: The permanent disability insurance.

In the next diagram on Figure 0.2 we allow transition back from "SICK" to "HEALTHY".

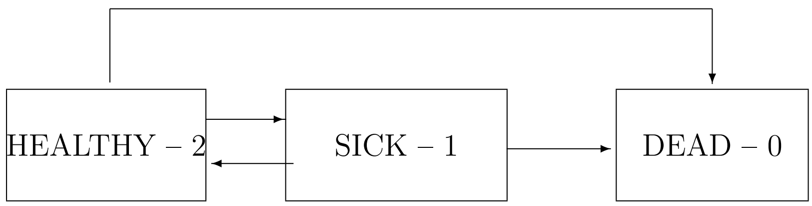

Figure 0.2: The disability income insurance model

In the next example we come back to the so called reversionary annuity. This is a good example to illustrate multi-state insurances, and details will be given in the next chapter.

Example 0.3 Consider the insurance from Example IV.4.3. In the present setting we do not assume the independence between lifes (x) and (y). We distinguish four states "HUSBAND ALIVE & WIFE ALIVE" (**), "HUSBAND ALIVE & WIFE DEAD"– $(*, \dagger)$ "HUSBAND DEAD & WIFE ALIVE" – $(\dagger, *)$ and "HUSBAND DEAD & WIFE DEAD" – $(*, *)$.

Furtheron we will use the following nomenclature¿ For HEALTHY we will use ACTIVE – A, for DISABLED we will write INVALID – I, and DEAD – D. Therefore we call our models AID; we have AID with recovery or without recovery.

[3]Convenient abbreviations: Healthy – Active – *, Disabled – Invalid – $\Diamond$, Dead - †

1 Markovian evolution of life

It turns out that a natural mathematical tool to model an evolution of states are Markov processes. Since this is a general theory we skip x in notation of the sequence or process. Thus in discrete case, $\mathcal{J}_j$ $(j = 0, 1, \ldots)$ is a nonhomegenous Markov chain with finite state space (DTMC), and for $\mathcal{J}(t)$ we use a nonhomogenous continuous time Markov chains (CTMC). These processes characterize an important memoryless property. In general it means that the conditional distribution of future evolution $\mathcal{F}_{t<}$ after time t, depends only on the present state $(\mathcal{F}_{=t})$, and not on a sequence of events that preceded it $(\mathcal{F}_{<t})$. That is

$$\Pr(\mathcal{F}_{t<}|\mathcal{F}_{=t} \vee \mathcal{F}_{<t}).$$

For DTMCs the above will be precisely stated in Proposition 1.2.

1.1 States evolves as a Markov chain

To begin with we recall a notion of a homogeneous Markov chain with state space $\mathcal{E} = \{0, 1, \ldots, m\}$. The basic notion is a *stochastic matrix* , which is a matrix $\mathbf{P} = (P^{ij})_{i,j=0}^m$, with nonegative elements, such that $\sum_{j=0}^m P^{ij} = 1$ for all $i \in \mathcal{E}$. We will use for convenience a matrix notations. If applied to a life (x) with many states, we can use notation $\mathbf{P}_x = (P_x^{ij})_{i,j=0}^m$. Thus we denote (i, j)-th element by $(\mathbf{P})^{ij} = P^{ij}$.

It is said that a sequence of $\mathcal{E}$-valued random variables X_k, $k = 0, 1, \ldots$ is a *nonhomogeneous Markov chain* defined by a family of *transition matrices* $\mathbf{P}(k) = (P^{ij}(k))_{ij=0}^{\infty}$, $k = 0, 1, \ldots$, if for $k = 1, 2, \ldots$, and $j_l \in \mathcal{E}$

$$\Pr(\mathcal{J}_1 = j_1, \ldots, \mathcal{J}_k = j_k|\mathcal{J}_0 = j_0) = P^{j_0 j_1}(0)P^{j_1 j_2}(1) \ldots P^{j_{k-1} j_k}(k-1). \quad (1.4)$$

Remark In the case when $\mathbf{P}(0) = \mathbf{P}(1) = \ldots$, we have a homogeneous DTMC. This is a frequently considered case, however in the presented theory is useless.

We make the following convention about order of indices of transition matrices:

Without proofs we state two elementary facts:

Proposition 1.1

$$\begin{aligned}\Pr(\mathcal{J}_k = j|\mathcal{J}_0 = i) &= \sum_{j_1,\dots,j_{k-1}\in\mathcal{E}} P^{ij_1}(0)\dots P^{j_{k-1}j}(k-1)\\ &= (\mathbf{P}(0)\mathbf{P}(1)\dots\mathbf{P}(k-1))^{ij}.\end{aligned}$$

Proposition 1.2 *A sequence of random variables is a DTMC defined by a family of stochastic matrices* $\mathbf{P}(k)$*,* $k = 0, 1, \dots$ *if and only if for all* $i, j, i_k \in \mathcal{E}$ *and* $k = 1, 2, \dots$

$$\begin{aligned}&\Pr(\mathcal{J}_k = j|\mathcal{J}_0 = j_0, \mathcal{J}_1 = j_1, \dots, \mathcal{J}_{k-1} = i)\\ &= \Pr(\mathcal{J}_k = j|\mathcal{J}_{k-1} = i) = P^{ij}(k-1).\end{aligned} \tag{1.5}$$

provided

$$\Pr(\mathcal{J}_0 = j_0, \mathcal{J}_1 = j_1, \dots, \mathcal{J}_{k-1} = i) > 0\ . \tag{1.6}$$

For $k < k'$ we define

$$\mathbf{P}(k, k') = \mathbf{P}(k)\mathbf{P}(k+1)\dots\mathbf{P}(k'-1).$$

Notice that $\Pr(\mathcal{J}_{k'} = j|\mathcal{J}_k = i) = (\mathbf{P}(k)\mathbf{P}(k+1)\dots\mathbf{P}(k'-1))^{ij}$.

In the continuous time there is a corresponding notion to $\mathbf{P}(k, k')$, which plays the role in definitions.

Example 1.3 We show that sequence of random variables $\mathcal{J}_{[x]+k}$, $k = 0, 1, \dots$ defined by (0.2) in Example 0.1 is a DTMC and explain the role of hypothesis HA. Notice that only sequences $j_0, \dots, j_{k-2}, i$ which are nonincreasing are of interest, because otherwise the condition (1.6) is not fulfilled. Thus, in the case $i = 1$

$$\begin{aligned}P_x^{11}(k-1) &= \Pr(\mathcal{J}_{[x]+k} = 1|\mathcal{J}_x = 1, \dots, \mathcal{J}_{[x]+k-1} = 1)\\ &= \Pr(\mathcal{J}_{[x]+k} = 1|\mathcal{J}_{[x]+k-1} = 1) = p_{[x]+k-1}\ ,\end{aligned}$$

and

$$\begin{aligned}P_x^{10}(k-1) &= \Pr(\mathcal{J}_{[x]+k} = 0|\mathcal{J}_x(0) = 1, \dots, \mathcal{J}_{[x]+k-1} = 1)\\ &= \Pr(\mathcal{J}_{[x]+k} = 0|\mathcal{J}_{[x]+k-1} = 1) = q_{[x]+k-1}\ .\end{aligned}$$

In case when $i = 0$

$$\begin{aligned} P_x^{01}(k-1) &= \Pr(\mathcal{J}_x(k) = 1 | \mathcal{J}_x(0) = 1, \ldots, \mathcal{J}_x(k-1) = 0) \\ &= \Pr(\mathcal{J}_{[x]+k} = 1 | \mathcal{J}_{[x]+k-1} = 0) = 0 \end{aligned}$$

and

$$\begin{aligned} P_x^{00}(k-1) &= \Pr(\mathcal{J}_{[x]+k} = 0 | \mathcal{J}_x = 1, \ldots, \mathcal{J}_{[x]+k-1} = 1) \\ &= \Pr(\mathcal{J}_{[x]+k} = 1 | \mathcal{J}_{[x]+k-1} = 0) = 1. \end{aligned}$$

Hence we have

$$\mathbf{P}(k-1) = \begin{pmatrix} 1 & 0 \\ q_{[x]+k-1} & p_{[x]+k-1} \end{pmatrix} .$$

Notice that if HA holds, then $\mathbf{P}_x(k) = \mathbf{P}(x+k)$, where $\mathbf{P}(0) = \mathbf{P}_0(0)$.

2 Continuous time Markov chains

2.1 Definition and basic properties

We say that $\mathcal{J}(t)$, $t \geq 0$ is a CTMC with state space $\mathcal{E}$, if for any sequence of instances $0 = t_0 < t_1 < \ldots$, sequence of random variables $X_k = \mathcal{J}(t_k)$, $k = 0, 1, \ldots$ is a DTMC. We often consider CTMC's with time parameter $0 \leq t \leq n$ and in the definition above we must write then $0 \leq t_0 < t_1 < \ldots \leq n$.

Consider a family of stochastic matrices $\mathbf{P}(t, t') = (P^{ij}(t, t'))_{i,j \in \mathcal{E}}$, where $0 \leq t \leq t'$, fulfilling

- $\mathbf{P}(t, t) = \mathbf{I}$ for all $t \geq 0$,
- for all $0 \leq t \leq s \leq t'$,

$$\mathbf{P}(t, t') = \mathbf{P}(t, s)\mathbf{P}(s, t') . \tag{2.7}$$

Such the family of matrices $\mathbf{P}(t, t')$, $0 \leq t \leq t'$ is said to be *transition matrix function* and (2.7) is called the *Chapman–Kolmogorov equation.*

Remark If $\mathbf{P}(0, y) = \mathbf{P}(t, t + y)$, for all $t, y \geq 0$, then the DTMC is called homogeneous. These are unrealistic chains for life insurance and therefore from now on by DTMC we mean a nonhomogeneous one.

Proposition 2.1 *Stochastic process $\mathcal{J}(t)$, $t \geq 0$ is a CTMC if and only if there exists a transition matrix function $\mathbf{P}(t, t')$, $0 \leq t \leq t'$ such that*

$$\begin{aligned}&\Pr(\mathcal{J}(t_1) = i_1, \ldots, \mathcal{J}(t_k) = i_k | \mathcal{J}(t_0) = i_0) \\ &= P^{i_0 i_1}(t_0, t_1) P^{i_1 i_2}(t_1, t_2) \ldots P^{i_{k-1} i_k}(t_{k-1}, t_k)\,, \end{aligned} \tag{2.8}$$

for $k = 0, 1, \ldots,\ i_0, i_1, \ldots, i_k \in \mathcal{E},\ 0 \leq t_0 \leq \ldots \leq t_k$.

We leave to the reader to prove this proposition. An equivalent form, which gives also an idea what is a Markov property, that is the *lack of the memory* of the past is stated in the following proposition. More precisely a process is a CTMC, if for any $t > 0$, its evolution after time t conditioned on the evolution prior t, depends only on the state of the process at t. A mathematical formulation of this sentence is left as Exercise 7.1.

In the following example we analyse process $\mathcal{J}_x(t)$ defined in the model ALIVE–DEAD for life (x). We will need here and later the following definition. It is said that a function $f : [0, n] \to \mathbb{R}$ fulfills condtion piecewise-continuous (abbreviated as PC), if for each $n' < n$, function f is continuous in $[0, n']$ apart from a finite number of point, and bounded. It can converges to infinity only at the end of the interval at n (finite or infinite).

Example 2.2 Suppose $T = T_0$ is the future lifetime of a new born person and define $\mathcal{J}(t)$ as in (0.1) of Example 0.1. Assume the the hypothesis HHP is valid, that is $\mu_x(t) = \mu(x + t)$, and that $\Pr(T > t) = \exp\left(-\int_0^t \mu(s)\,\mathrm{d}s\right)$. We show that $\mathcal{J}(t)$ is a CTMC. Namely for $k \geq 1,\ i_0, i_1, \ldots, i_n \in \{0, 1\}$ i $t_0 < t_1 < \ldots < t_k$,

$$\begin{aligned}&\Pr(\mathcal{J}(t_k) = i_k \mid \mathcal{J}(t_{k-1}) = i_{k-1}, \ldots, \mathcal{J}(t_1) = i_1, \mathcal{J}(t_0) = i_0) \\ &= \Pr(\mathcal{J}(t_k) = i_k \mid \mathcal{J}(t_{k-1}) = i_{k-1})\,, \end{aligned} \tag{2.9}$$

provided $\Pr(\mathcal{J}(t_{n-1}) = i_{n-1}, \ldots, \mathcal{J}(t_1) = i_1, \mathcal{J}(t_0) = i_0) > 0$ (this means that $j_0, j_\cdot, \ldots$ is a nonincreasing sequence of ones and zeros). Next

$$\Pr(\mathcal{J}(t_n) = i_n \mid \mathcal{J}(t_{n-1}) = i_{n-1}) = \begin{cases} {}_{t_k - t_{k-1}}p_{t_{k-1}} & i_k = 1 \\ {}_{t_k - t_{k-1}}q_{t_{k-1}} & i_k = 0 \end{cases}$$

Hence

$$\mathbf{P}(t,t') = \begin{pmatrix} 1 & 0 \\ {}_{t'-t}q_t & {}_{t'-t}p_t \end{pmatrix} .$$

Consider now

$$\begin{aligned} {}_{t'-t}q_t = P^{\mathrm{AD}}(t,t') &= \Pr(T_0 \le t' \mid T_0 > t) && (2.10) \\ &= 1 - \exp\Big(-\int_t^{t'} \mu(s)\,\mathrm{d}s\Big) . && (2.11) \end{aligned}$$

for $t \le t'$. Hence

$${}_{t'-t}q_t = \int_t^{t'} \mu(s)\,\mathrm{d}s + o(\int_t^{t'} \mu(s)\,\mathrm{d}s).$$

Suppose now that the mortality rate $\mu(t)$ fulfills condition PC. (we do not exclud infinity at the end of $[0,n]$). Let $0 < t < n$ be such that $\mu(t-) = \mu(t+)$. Then, for $t' \downarrow t$

$$\int_t^{t'} \mu(s)\,\mathrm{d}s = \mu(t+) + o(t'-t)$$

and for $t' \uparrow t$

$$\int_{t'}^{t} \mu(s)\,\mathrm{d}s = \mu(t-) + o(t-t').$$

Similarly we can demonstrate $P^{\mathrm{DD}}(t,t') = P^{\mathrm{DA}}(t,t') = 0$. Hence

$$\lim_{t'\downarrow} \frac{\mathbf{P}(t,t') - \mathbf{I}}{t'-t} = \boldsymbol{\mu}(t) = \begin{pmatrix} 0 & 0 \\ \mu(t) & -\mu(t) \end{pmatrix} ,$$

for all t being a continuity point of $\mu(t)$. For other points we simply set $\boldsymbol{\mu} = \mathbf{0}$. Notice the row corresponding to absorbing state 0 has all zero entries.

In this example we see, that if the distribution F of T has atoms, then we have problems with definition of intensity matrix $\boldsymbol{\mu}$.

We now introduce conditions when the transition matrix function exists:

[Z1] Function $\mathbf{P}(t,t')$ is continuous in t' for all t and is continuous in t for all t', ie. $\lim_{t'\downarrow 0}\mathbf{P}(0,t')=\mathbf{I}$ and

$$\lim_{t'\downarrow t}\mathbf{P}(t,t')=\lim_{t'\uparrow t}\mathbf{P}(t',t)=\mathbf{I}\,. \tag{2.12}$$

[Z2] Outside a finite set $\mathcal{N}\subset[0,n]$, there exists

$$\boldsymbol{\mu}(t)=\lim_{t'\downarrow t}\frac{\mathbf{P}(t,t')-\mathbf{I}}{t'-t}=\lim_{t'\uparrow t}\frac{\mathbf{P}(t',t)-\mathbf{I}}{t-t'},\qquad t\notin\mathcal{N}\,. \tag{2.13}$$

For completeness, we define at points $t\in\mathcal{N}$ values of $\boldsymbol{\mu}(t)$ such that μ^{ij} are right continuous with lefthand limits.

[Z3] Entries of $\boldsymbol{\mu}(t)$ fulfill condition PC.

Proposition 2.3 *For $i\neq j$, $\mu^{ij}(t)\geq 0$, $\mu^{ii}(t)\leq 0$ and for $i\in\mathcal{E}$ oraz $t\geq 0$,*

$$\sum_{j\in\mathcal{E}}\mu^{ij}(t)=0 \tag{2.14}$$

for all t. Moreover, if $i\in\mathcal{E}_{\mathrm{abs}}$ is absorbing, then $\mu^{ij}(t)=0$, $j\in\mathcal{E}$.

Next we set

$$\mu^i(t)=\sum_{j\neq i}\mu^{ij}(t).$$

Sometimes the following convention for the numeration seems to be convenient and will be used in the text:

$$\boldsymbol{\mu}(t)=\begin{pmatrix}-\mu^0(t) & \mu^{0\,1}(t) & \dots & \mu^{0\,m}(t)\\ \mu^{1\,0}(t) & -\mu^1(t) & \dots & \mu^{1\,m}(t)\\ \vdots & \vdots & \dots & \vdots\\ \mu^{m0}(t) & \mu^{m\,1}(t) & \dots & -\mu^m(t)\end{pmatrix}$$

Notice that in the language of Markov chain $\mathcal{J}(t)$ we have $\Pr(\mathcal{J}(t+h) \neq i|\mathcal{J}(t) = i) = h\mu^i(t) + o(h)$, provided $t \neq \mathcal{N}$. Similarly for$i \neq j$ we have $\Pr(\mathcal{J}(t+h) = j|\mathcal{J}(t) = i) = h\mu^{ij}(t) + o(h)$.

Matrix function $\boldsymbol{\mu}(t), t \geq 0$ is said to be an *transition intensities matrix* of CTMC $\mathcal{J}(t)$. Hypothesis HHJ means $\boldsymbol{\mu}_x(t) = \boldsymbol{\mu}(x+t)$.

4

Example 2.4 *Makeham intensity matrix* has entries for $i \neq j$ $\mu^{ij}(t) = A^{ij} + B^{ij}(c^{ij})^t$, where we postulate $B^{ij} \geq 0$, $A^{ij} \geq -B^{ij}$ i $c^{ij} > 1$.

Example 2.5 Consider process $\mathcal{J}_x(t)$ defined by 0.3,which describes state of life (x) in the multi-decrement model. Notice that here we have absorbing states $\mathrm{D}_1, \ldots, \mathrm{D}_m$ and one transiton state A. Notations and defintions of (T_x, J_x) are from Chapter IV and we assume that hypothesis HHP-D holds. Under some assumption on piecewise continuity of $f_x(t,j)$ we can show

$$\begin{pmatrix} \mu_x^{\mathrm{D}_1\mathrm{A}}(t) & \mu_x^{\mathrm{D}_1\mathrm{D}_1}(t) & \ldots & \mu_x^{\mathrm{D}_1\mathrm{D}_m}(t) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_x^{\mathrm{D}_m\mathrm{A}}(t) & \mu_x^{\mathrm{D}_m\mathrm{D}_1}(t) & \ldots & \mu_x^{\mathrm{D}_m\mathrm{D}_m}(t) \\ \mu^{AA}(t) & \mu_x^{\mathrm{A}\mathrm{D}_1}(t) & \ldots & \mu_x^{\mathrm{A}\mathrm{D}_m}(t) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \\ -\mu_x^{(\tau)}(t) & \mu_x^{(1)}(t) & \ldots & \mu_x^{(m)}(t) \end{pmatrix}. \tag{2.15}$$

Theorem 2.6 *Let $\mathcal{N}$ be the discontinuity set of $\boldsymbol{\mu}(t)$. For $t \neq \mathcal{N}$ there exist partial derivatives $\partial/(\partial t)\mathbf{P}^{ij}(t,t')$ and for $t' \neq \mathcal{N}$ there exist partial derivatives $\partial/(\partial t')\mathbf{P}(t,t')$. Matrix function $\mathbf{P}(t,t')$ $0 \leq t \leq t' \leq n$ fulfills:*

[4]Zaoenia [Z.i] pozwalaj na udowodnienie w teorii procesow stochastycznych, ze realizacje sa kawakami stałe, prawostronnie cige i z lewostronnymi granicami. do tego wrcimy w podrozdziale 2.2
(A.ii) dla $i \neq j$

$$\sup_{a \leq t \leq b, h>0} \frac{P^{ij}(t,t+h)}{h} < \infty.$$

[retrospective Kolmogorov differentional equation]

$$\frac{\partial}{\partial t}\mathbf{P}(t,t') = -\boldsymbol{\mu}(t)\mathbf{P}(t,t') \tag{2.16}$$

[prospective Kolmogorov differentional equation]

$$\frac{\partial}{\partial t'}\mathbf{P}(t,t') = \mathbf{P}(t,t')\boldsymbol{\mu}(t') \tag{2.17}$$

and the boundary condition is $\mathbf{P}(t,t) = \mathbf{I}$ *for* $t \geq 0$.

Proof Suppose $t \leq t+h \leq t'$

$$\begin{aligned}&\frac{1}{h}(\mathbf{P}(t+h,t') - \mathbf{P}(t,t') \\ &\quad = \frac{1}{h}\mathbf{P}(t+h,t') - \mathbf{P}(t,t+h)\mathbf{P}(t+h,t') \\ &\quad = \frac{1}{h}(\mathbf{I} - \mathbf{P}(t,t+h))\mathbf{P}(t+h,t').\end{aligned}$$

Corollary 2.7 *Suppose that* i *is a state such that* $\mu^{ij}(t) = 0$ *for all* $j \neq i$. *We have*

$$P^{ii}(t,t') = \exp(-\int_t^{t'} \mu^i(s)\,ds).$$

Proof In this case from the retrospective Kolmogorov equation (2.36) we obtain

$$\frac{\partial}{\partial t}P^{ii}(t,t') = -\mu^i(t)P^{ii}(t,t').$$

The proof is completed when solving this differential equation with initial condition $P^{ii}(t,t) = 1$. □

Example 2.8 For the model AD with mortality rate set $\mu^{\mathrm{AD}}(t) = \mu(t)$.

The prospective Kolmogorov equation are

$$\frac{\partial}{\partial t'} P^{\mathrm{AA}}(t,t') = -P^{\mathrm{AA}}(t,t')\mu(t) \tag{2.18}$$
$$\frac{\partial}{\partial t'} P^{\mathrm{AD}}(t,t') = P^{\mathrm{AA}}(t,t')\mu(t) \tag{2.19}$$
$$\frac{\partial}{\partial t'} P^{\mathrm{DA}}(t,t') = -P^{\mathrm{DA}}(t,t')\mu(t) \tag{2.20}$$
$$\frac{\partial}{\partial t'} P^{\mathrm{DD}}(t,t') = P^{\mathrm{DA}}(t,t')\mu(t') \,. \tag{2.21}$$

and the boundary conditions are

$$P^{\mathrm{AA}}(t,t) = 1 \tag{2.22}$$
$$P^{\mathrm{AD}}(t,t) = 0 \tag{2.23}$$
$$P^{\mathrm{DA}}(t,t) = 0 \tag{2.24}$$
$$P^{\mathrm{DD}}(t,t) = 1. \tag{2.25}$$

Solving the first equation we have

$$P^{\mathrm{AA}}(t,t') = e^{-\int_t^{t'} \mu(s)\,ds}.$$

Solving next the third equuation we have $P^{\mathrm{DA}}(t,t') = 0$. Hence from the fourth we have $P^{\mathrm{DD}}(t,t') = 1$. We leave the reader to demonstrate that

$$P^{\mathrm{AD}}(t,t') = 1 - e^{-\int_t^{t'} \mu(s)\,ds}.$$

Integrating out from t to t' (2.36) and (2.37) respectively, we can prove:

Theorem 2.9 *We have*

$$\mathbf{P}(t,t') = \mathbf{I} + \int_t^{t'} \boldsymbol{\mu}(s)\mathbf{P}(s,t')\,\mathrm{d}s \tag{2.26}$$

and

$$\mathbf{P}(t,t') = \mathbf{I} + \int_t^{t'} \mathbf{P}(t,s)\boldsymbol{\mu}(s)\,\mathrm{d}s\,. \tag{2.27}$$

for all $0 \le t < t'$.

One can show that there exists only one transition matrix function $\mathbf{P}(t,t')$ solving (2.26) and (2.27) respectively.

Theorem 2.10 *For* $0 \leq t \leq t'$,

$$\mathbf{P}(t,t') = \mathbf{I} + \sum_{n=1}^{\infty} \int_t^{t'} \int_{s_1}^{t'} \dots \int_{s_{n-1}}^{t'} \boldsymbol{\mu}(s_1) \dots \boldsymbol{\mu}(s_n)\, \mathrm{d}s_n \dots \mathrm{d}s_1 \tag{2.28}$$

or

$$\mathbf{P}(t,t') = \mathbf{I} + \sum_{n=1}^{\infty} \int_t^{t'} \int_t^{s_1} \dots \int_t^{s_{n-1}} \boldsymbol{\mu}(s_n) \dots \boldsymbol{\mu}(s_1)\, \mathrm{d}s_n \dots \mathrm{d}s_1 \,. \tag{2.29}$$

Proof Equations (2.28), and (2.29) are classical Volterra integral equations of type

$$\mathbf{x}(t) = \mathbf{I} + \int_t^{t'} \mathbf{x}(s)\mathbf{A}(s)\, ds \tag{2.30}$$

and

$$\mathbf{x}(t) = \mathbf{I} + \int_t^{t'} \mathbf{A}(s)\mathbf{x}(s)\, ds. \tag{2.31}$$

If we assume that $\mathbf{A}(s)$ are matrix function, such that $\int_t^{t'} ||A(s)||\, ds < \infty$, then equations 2.30 and 2.31 are solved by the so called Pickard's sum

$$\mathbf{x}(t) = \mathbf{I} + \sum_{j=1}^{\infty} I_j(x),$$

where

$$I_n(x) = \int_t^{t'} \int_{s_1}^{t'} \dots \int_{s_{n-1}}^{t'} \mathbf{A}(s_1) \dots \mathbf{A}(s_n)\, ds_n \dots ds_1.$$

Note that changing the order of integration we obtain

$$I_n(x) = \int_t^{t'} \mathbf{A}(s_1) \int_t^{s_1} \dots \int_t^{s_{n-1}} \mathbf{A}(s_n) \dots \mathbf{A}(s_n)\, ds_1 \dots ds_n.$$

One must prove the convergence

$$\sum_{j=1}^{\infty} I_j(x).$$

Details of the proofs for the convergence and the uniqueness can be found in Baake and Schlagel . [5]

[5]Michael Baake and Ulrike Schlagel The Peano-Baker Series arXiv:1011.1775v2 [math.CA] 6 Jan 2012

Example 2.11 Consider the case when the matrix intensity function is constant, i.e. $\boldsymbol{\mu}(t) \equiv \boldsymbol{\Lambda}$. Then from equation (2.28) we obtain

$$\mathbf{P}(t,t') = \sum_{l=0}^{\infty} \frac{(\boldsymbol{\Lambda}(t'-t))^l}{l!}.$$

This so called the *matrix exponential function*. Thus

$$\mathbf{P}(t,t') = e^{\boldsymbol{\Lambda}(t'-t)}.$$

The matrix exponential function has a property

$$e^{\boldsymbol{\Lambda}(t'-t)} = (e^{\boldsymbol{\Lambda}})^{(t'-t)}.$$

Hence for dla $k = 0, \ldots$ and $0 \leq u \leq 1$ we have

$$\mathbf{P}(k,k+u) = (\mathbf{P}(k,k+u))^u \ . \qquad (2.32)$$

Example 2.12 Suppose $\boldsymbol{\mu}(t) = \boldsymbol{\mu}(\lfloor t \rfloor)$, for all t. This condition correponds to classical HCFM hypothesis for one life and one risk; see Chapter I.2.1. Then we have

$$\mathbf{P}(x,x+n+u) \quad = \quad \mathbf{P}(x,x+n)(\mathbf{P}(x+n,x+n+1))^u \ . \qquad (2.33)$$

Furtheron

$$\mathbf{P}(x,x+n) = \exp(-\sum_{k=0}^{n-1} \boldsymbol{\mu}(x+k)) \ . \qquad (2.34)$$

To demonstrate this fact, consider $\mathbf{P}(k,k+u)$ for some $k = 0,1,\ldots$. Clearly we have

$$\mathbf{P}(x,x+n+u) = \mathbf{P}(x,x+n)(\mathbf{P}(x+n,x+n+u)) \ .$$

Using example 2.11 we obtain formula (2.33). We leave to the reader to demonstrate formula (2.34).

2.2 What can we say about evolution

Consider a CTMC $\mathcal{J}(t)$ with intensity matrix $\boldsymbol{\mu}(t)$. We will show an algorythm for an evolution of $\mathcal{J}(t)$ when $t \geq 0$. Assume state space is $\mathcal{E} = \{0,1,\ldots,m\}$ and $\mathcal{J}(0) = i_0$. In actuarial practice the initial state is state 'HEALTHY' or 'Active'. We consider the theory from Chapter IV on one life with few risks and say that the status is 'being in state i_0. Forces of decrement are $\mu^{i_0 0}(t), \ldots \mu^{i_0 m}(t)$ (without $\mu^{i_0 i_0}(t)$), where $i_0 \in \mathcal{E}_{\rm tra}$. Denote time to exit from the status because of risk J_0 by T_0. Random vector (T_0, J_0) has the density function $f_0(t, i_1) = \mu^{i_0 i_1}(t) \exp[-\int_0^t \mu^{i_0}(s)\ {\rm d}s]$.

Suppose now $\mathcal{J}(t) = i$. Since the process is a CTMC, the past before t is not important. The status now is i and time to exit from the status is $T_{t,i}$ and let $J_{t,i}$ denotes the corresponding decrement, that is instant of the next jump after t is $t + T_{t,i}$ and the process jumps to $J_{t,i} = j$. Again we can write the density function of $(T_{t,i}, J_{t,i})$ $f_0(t, i_1) = \mu^{ij}(t) \exp[-\int_0^t \mu^i(s)\ {\rm d}s]$.

A standard notaton in the multi-state theory is

$$P^{\underline{ii}}(t, t') = \Pr(\text{process is in state } i \text{ in interval } [t, t'] | \mathcal{J}(t) = i).$$

Thus $P^{\underline{ii}}(0, t) = \Pr(T_0 > t)$. Notice that unless $\mu^{ij}(t) = 0$ for $i \neq j$ we have $P^{\underline{ii}}(0, t) \neq P^{ii}(0, t)$.

Using these facts we can design a simulation algorthm to generate a trajectory of $\mathcal{J}(t)$. Thus if a jump is at instant t, then we draw the next jump according the density function

$$\mu^i(s) \exp[-\int_t^{t+s} \mu^i(v)\ {\rm d}v], \qquad s > 0. \tag{2.35}$$

or the tail distribution function

$$\Pr(T_{t,i} > s) = \exp[-\int_t^{t+s} \mu^i(v)\ {\rm d}v], \qquad s > 0.$$

Thus if U is a uniformly distributed r.v. on $[0, 1]$, then the moment of the next jump is the solution of

$$-\log U = \int_t^{t+s} \mu^i(v)\, dv$$

with respect variable s. If the result is s, then we draw the state where the process is jumping to j with probability

$$\frac{\mu^{ij}(t+s)}{\mu^i(t+s)}, j \neq i.$$

We have to remember, that between jumps the process is constant.

Remark From the algorithm above we can see that because functions $\mu^{ij}(t)$ are locally bounded, there can be a finite number of jumps in finite intervals.

Exercises; on line lecture 6

1. Show that transition functions fulfill the following system of partial diferential equation: for $t \notin \mathcal{N}$

$$\frac{\partial}{\partial t}P^{ij}(t,t') = \mu^{i}(t)P^{ii}(t) - \sum_{k\neq i}\mu^{ik}(t)P^{kj}(t,t'), \qquad t \neq \mathcal{N} \tag{2.36}$$

and for $t' \notin \mathcal{N}$

$$\frac{\partial}{\partial t'}P^{ij}(t,t') = -P^{ij}(t)\mu^{j}(t) + \sum_{k\neq i}P^{ik}(t,t')\mu^{kj}(t') \qquad t' \neq \mathcal{N}. \tag{2.37}$$

2. Write the matrix transition intensities matrix for AID models (with recoveries and without). Write prospective and retrospective Kolmogorov equations for AID without recoveries.

3. For a group of m lifes with independent lifetimes $T_1, \ldots, T_m$ let $\mathcal{J}(t)$ be the process counting the number of alives at time t, ie.

$$\mathcal{J}(t) = \begin{cases} m & \text{jeli } 0 \leq t < T_{(1)} \\ m-1 & \text{jeli } T_{(1)} \leq t < T_{(2)} \\ \vdots & \vdots \\ 1 & \text{jeli } T_{(m-1)} \leq t < T_{(m)} \\ 0 & \text{jeli } T_{(m)} \leq t, \end{cases} \tag{2.38}$$

as usual $(T_{(1)}, \ldots, T_{(m)})$ denotes the order statistics of variables $(T_1, \ldots, T_m)$. Show that if $(T_1, \ldots, T_m)$ are independent identically distributed with mortality rate $\mu(t)$, then $\mathcal{J}(t)$ is a CTMC. Find the matrix of transition intensities.