Survival Analysis, Master Equation, Efficient Simulation of Path-Related Quantities, and Hidden State Concept of Transitions

Dirk Helbing

II. Institute of Theoretical Physics, Pfaffenwaldring 57/III, 70550 Stuttgart, Germany

1   Introduction

This paper presents and derives the interrelations between survival analysis and master equation. Both have important applications in the social sciences and other scientific fields treating *stochastic systems*. However, since they focus on different aspects of modeling, it is not yet generally known that they are closely related to each other.

Survival analysis deals with modeling the transitions between succeeding states of a system. Questions related with this are the *timing, spacing,* and *sequencing* of the states of a time series. Survival analysis tries to fit and understand the distribution of these quantities in terms of the functional form of the *hazard rates* which are responsible for the investigated transitions. The parameters specifying the concrete functional form of the hazard rates are normally estimated from empirical data by means of the *maximum-* or *partial-likelihood method.*

Once the hazard rates are empirically known or hypothetically specified, one can carry out microsimulations of corresponding time series by means of the *Monte-Carlo simulation method.* This allows the prognosis and investigation of the characteristics and frequency of time series.

However, if one is interested in *cross-sectional data* connected with the stochastic process under consideration, one needs to know the temporal evolution of the *distribution of states.* This can be obtained by simulation of the associated *master equation,* which only presupposes that the initial distribution and the hazard rates are given. In other words: The master equation is suitable for the calculation and prognosis of *cross-sectional data* which are related with the longitudinal life-table data used for survival analysis.

In addition, some new formulas are introduced which allow the determination of path-related (i.e. longitudinal) quantities like the *occurence probability*, the *occurence time distribution*, or the *effective cumulative life-time distribution* of a certain sequencing of states *(path)*. These can be efficiently evaluated with a recently developed simulation tool (EPIS) which also provides a new solution method of the master equation (the *contracted path integral solution*). In contrast, a calculation on the basis of time-series data would require an extreme computational effort.

The effective cumulative life-time distribution facilitates the formulation of a *hidden state concept of behavioral changes* which allows an interpretation of the respective

time-dependence of hazard rates. Hidden states represent states which are either not phenomenological distinguishable from other states, not externally measurable, or simply not detected. They could, for example, reflect the psychological and mental stages preceding a concrete action, the individual predisposition, or the internal attitude towards a certain behavior.

2 Survival Analysis

2.1 Preliminaries

Survival analysis deals with the investigation of the *timing, spacing,* and *sequencing* of *longitudinal time series (panel studies),* e.g. life-table data (Blossfeld et al. 1986; Kalbfleisch/Prentice 1980; Elandt-Johnson/Johnson 1980; Diekmann/Mitter 1984; Tuma/Hannan 1984; Cox/Oakes 1984; Lancaster 1990; Courgeau/Lelièvre 1992). These time series have the general form

$$Y(t) = Y_n \qquad \text{for} \qquad T_n \leq t < T_{n+1} \,, \tag{1}$$

where T_n is the variable of the *waiting time* after which the *state Y_n* is occupied. The state Y_n remains occupied for a time period $T_n \leq t < T_{n+1}$ ending directly before the next *event* (Y_{n+1}, T_{n+1}). Two succeeding events (Y_{n-1}, T_{n-1}) and (Y_n, T_n) determine the nth *episode* $((Y_{n-1}, T_{n-1}), (Y_n, T_n))$ of a time series.

The state Y_n stems out of a *set $\mathcal{M}_n$ of possible outcomes* which are often called *risks*. Without loss of generality[1] we can assume that this set does not depend on the respective episode n which implies $\mathcal{M}_n = \mathcal{M}$. Furthermore, $\mathcal{M}$ is normally a *discrete* set so that

$$Y_n \in \mathcal{M} = \{1, 2, \ldots, i, \ldots, N\} \,. \tag{2}$$

The elements i of this set represent the possible outcomes. Examples for $\mathcal{M}$ are:

$$\mathcal{M} = \{\text{single, married, widowed, divorced}\} \tag{3}$$

or

$$\mathcal{M} = \{\text{school, army, training, studies, employment, unemployment, retirement}\} \,. \tag{4}$$

Of course, the sequence of waiting times $T_n = t_n$ *('timing')* and often also the sequence of states $Y_n = i_n$ *('sequencing')* vary individually so that the actual time series

$$Y_\alpha(t) = i_n \qquad \text{for} \qquad t_n \leq t < t_{n+1} \tag{5}$$

can only be described by a *stochastic process*

$$(Y, T) := \{(Y_n, T_n) : n = 1, 2, \ldots\} \,. \tag{6}$$

The subscript $\alpha \in \{1, 2, \ldots, A\}$ distinguishes the different individuals, systems, or realizations to which the respective time series belong.

---

[1]One can simply define $\mathcal{M} := \bigcup \mathcal{M}_n$, where "$\bigcup$" symbolizes the union of sets. If, for a given event history H_{n-1}, no transitions take place to state $j \in \mathcal{M}$ one must only set the corresponding transition rate equal to zero, e.g. for $j = i_{n-1}$.

Now, the *spacing* of a time series can be defined as the sequence of *life times (survival times, failure times)*

$$V_n := T_n - T_{n-1}, \qquad k = 1, 2, \ldots. \tag{7}$$

Additionally, we introduce the so-called *event history* of a state $Y_n = i_n$ which has been occupied at time $T_n = t_n$ by

$$H_{n-1} := \{t_0, i_0; t_1, i_1, \mathbf{x}_1; \ldots; t_{n-1}, i_{n-1}, \mathbf{x}_{n-1}\}. \tag{8}$$

Here, the *vector of covariates* $\mathbf{x}_k = (x_{k,1}, \ldots, x_{k,M})$ comprises the different factors $x_{k,m}$ like education or sex which may influence the *transition* (i.e. change) from the $(k-1)$st state $Y_{k-1} = i_{k-1}$ to the kth state $Y_k = i_k$.

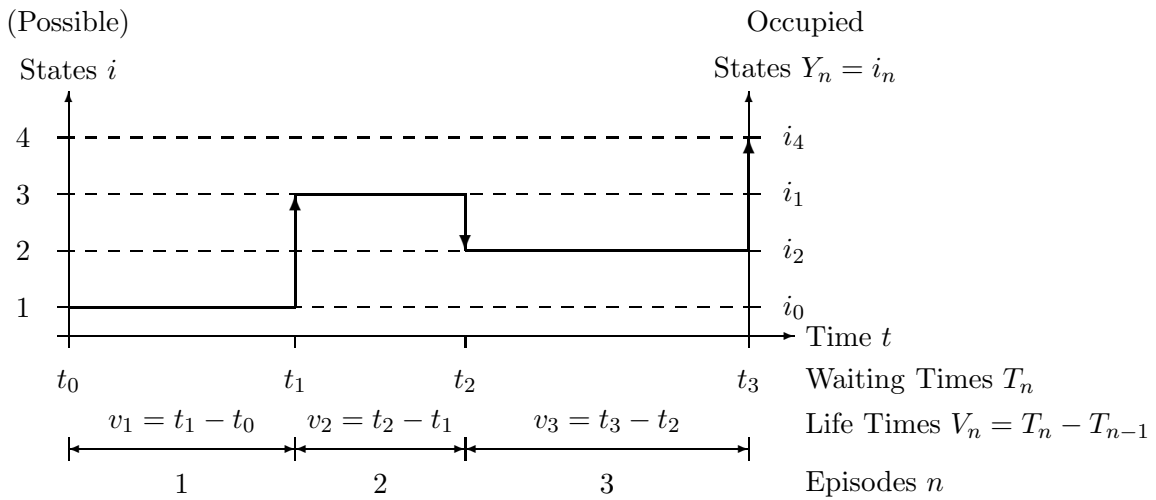The following figure illustrates the terms which were introduced above:

Figure 1: Example illustrating a time series with $N = 4$ competing risks and 3 episodes which are given by the sequence of events $(1, t_0) \to (3, t_1) \to (2, t_2) \to (4, t_3)$.

2.2 Central Concepts

We now come to the definition of some central concepts of survival analysis. In view of the following discussion, we restrict ourselves to the special case that

1. only the last state $Y_{n-1} = i_{n-1}$ of the event history H_{n-1} has an influence on the nth transition *(Markov case)*,

2. the transitions and the vector of covariates $\mathbf{x}_n$ are independent of the respective episode n.

The second assumption implies that the vector of covariates $\mathbf{x}$ is time-independent during the survey. This means that each individual or system α is characterized by a fixed value of $\mathbf{x}$ and that different vectors of covariates distinguish different *cohorts (subpopulations)*.

Then, the so-called *hazard rates (transition rates)* $\lambda_{\mathbf{x}}(t; j|i)$ of subpopulation $\mathbf{x}$ are defined as the probability $P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j | T_n \geq t, Y_{n-1} = i)$[2] per unit time $\Delta t > 0$

---

[2]Instead of $P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j | T_n \geq t, Y_{n-1} = i)$ one sometimes writes $P_{\mathbf{x}}(t \leq T_n < t + \Delta t, Y_n = j | T_n \geq t, Y_{n-1} = i)$. However, the expression $t \leq T_n < t + \Delta t$ is unnecessary complicated since $t \leq T_n$ is already presupposed by the condition $T_n \geq t$ in the second part of the argument of $P_{\mathbf{x}}$.

to change into state $Y_n = j$ up to time $t + \Delta t$ on the conditions that the nth transition did not happen before time t and the preceding state was $Y_{n-1} = i$:

$$\lambda_{\mathbf{x}}(t; j|i) := \lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j | T_n \geq t, Y_{n-1} = i). \tag{9}$$

Moreover, the *survivor function (survival function)* $S_{\mathbf{x}}^n(t|i)$ of cohort $\mathbf{x}$ is defined by the probability $P_{\mathbf{x}}(T_n \geq t | Y_{n-1} = i)$ that the nth transition does not take place before time T_n, given that the preceding state was $Y_{n-1} = i$:

$$S_{\mathbf{x}}^n(t|i) := P_{\mathbf{x}}(T_n \geq t | Y_{n-1} = i). \tag{10}$$

One can find the relation

$$S_{\mathbf{x}}^n(t|i) = 1 - F_{\mathbf{x}}^n(t|i) \tag{11}$$

with regard to the *cumulative distribution function (life-time distribution, failure distribution, duration distribution function)*

$$F_{\mathbf{x}}^n(t|i) := P_{\mathbf{x}}(T_n < t | Y_{n-1} = i) \tag{12}$$

of subpopulation $\mathbf{x}$. The latter describes the probability that the nth transition happens before time t, given that the preceding state was $Y_{n-1} = i$. If we are interested in the probability that the nth transition not only happens before time t but additionally leads to state $Y_n = j \neq i$, we need the quantity

$$F_{\mathbf{x}}^n(t; j|i) := P_{\mathbf{x}}(T_n < t, Y_n = j | Y_{n-1} = i) \tag{13}$$

which fulfils

$$\sum_{\substack{j=1 \\ (j \neq i)}}^{N} F_{\mathbf{x}}^n(t; j|i) = P_{\mathbf{x}}(T_n < t | Y_{n-1} = i) = F_{\mathbf{x}}^n(t|i). \tag{14}$$

Finally we define the *probability density function (failure time (sub)density function for failure type j)* by

$$
\begin{aligned}
f_{\mathbf{x}}^n(t; j|i) \ &:= \ \lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(t \leq T_n < t + \Delta t, Y_n = j | Y_{n-1} = i) \\
&= \ \lim_{\Delta t \to 0} \frac{P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j | Y_{n-1} = i) - P_{\mathbf{x}}(T_n < t, Y_n = j | Y_{n-1} = i)}{\Delta t} \\
&= \ \frac{d}{dt} F_{\mathbf{x}}^n(t; j|i).
\end{aligned}
\tag{15}
$$

(9) to (11), and (15) imply the relations

$$\sum_{\substack{j=1 \\ (j \neq i)}}^{N} f_{\mathbf{x}}^n(t; j|i) = \frac{d}{dt} F_{\mathbf{x}}^n(t|i) = -\frac{d}{dt} S_{\mathbf{x}}^n(t|i) \tag{16}$$

and

$$
\begin{aligned}
\lambda_{\mathbf{x}}(t; j|i) S_{\mathbf{x}}^n(t|i) &= \lambda_{\mathbf{x}}(t; j|i) P_{\mathbf{x}}(T_n \geq t | Y_{n-1} = i) \\
&= \ \lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j | T_n \geq t, Y_{n-1} = i) P_{\mathbf{x}}(T_n \geq t | Y_{n-1} = i) \\
&= \ \lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j, T_n \geq t | Y_{n-1} = i) \\
&= \ f_{\mathbf{x}}^n(t; j|i).
\end{aligned}
\tag{17}
$$

Equations (16) and (17) yield the differential equation

$$\sum_{\substack{j=1 \\ (j \neq i)}}^{N} \lambda_{\mathbf{x}}(t; j|i) S_{\mathbf{x}}^{n}(t|i) = -\frac{d}{dt} S_{\mathbf{x}}^{n}(t|i) \tag{18}$$

which has the important solution

$$S_{\mathbf{x}}^{n}(t|i) = \exp\left[-\int_{t_{n-1}}^{t} dt' \, \lambda_{\mathbf{x}}(t'|i)\right] = \prod_{\substack{j=1 \\ (j \neq i)}}^{N} \exp\left[-\int_{t_{n-1}}^{t} dt' \, \lambda_{\mathbf{x}}(t'; j|i)\right] \tag{19}$$

with the *overall departure rate (overall hazard rate, overall failure rate)*

$$\lambda_{\mathbf{x}}(t|i) := \sum_{\substack{j=1 \\ (j \neq i)}}^{N} \lambda_{\mathbf{x}}(t; j|i) \,. \tag{20}$$

That is, survivor functions are determined by exponential relations which only depend on the hazard rates.

3 The Master Equation Technique

3.1 Derivation of the Master Equation

We will now derive a system of differential equations for the evolution of the *probability distribution* $P_{\mathbf{x}}(j, t)$ of states j with time t which is associated with the above described stochastic process. (Each subpopulation $\mathbf{x}$ obeys its own system of equations.) For this purpose we apply two relations from probability theory:

$$\sum_{i=1}^{N} P_{\mathbf{x}}(i, t'|j, t) = 1 \tag{21}$$

and

$$P_{\mathbf{x}}(j, t') = \sum_{i=1}^{N} P_{\mathbf{x}}(j, t'|i, t) P_{\mathbf{x}}(i, t) \,. \tag{22}$$

Here, $P_{\mathbf{x}}(j, t'|i, t)$ denotes the probability that we have state j at time t' given that we had state i at time t. With (21) and (22) we get

$$\begin{aligned}
&\lim_{\Delta t \to 0} \frac{P_{\mathbf{x}}(j, t + \Delta t) - P_{\mathbf{x}}(j, t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\left(\sum_{i=1}^{N} P_{\mathbf{x}}(j, t + \Delta t|i, t) P_{\mathbf{x}}(i, t)\right) - \left(\sum_{i=1}^{N} P_{\mathbf{x}}(i, t + \Delta t|j, t)\right) P_{\mathbf{x}}(j, t)\right] \\
&= \sum_{\substack{i=1 \\ (i \neq j)}}^{N} \left[\left(\lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(j, t + \Delta t|i, t)\right) P_{\mathbf{x}}(i, t) - \left(\lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(i, t + \Delta t|j, t)\right) P_{\mathbf{x}}(j, t)\right] \,.
\end{aligned} \tag{23}$$

Due to

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(j, t + \Delta t|i, t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P_{\mathbf{x}}(T_n < t + \Delta t, Y_n = j|T_n \geq t, Y_{n-1} = i) \tag{24}$$

and (9) we finally obtain the differential equation

$$\frac{d}{dt}P_{\mathbf{x}}(j,t) = \sum_{\substack{i=1 \\ (i \neq j)}}^{N} \left[\lambda_{\mathbf{x}}(t;j|i)P_{\mathbf{x}}(i,t) - \lambda_{\mathbf{x}}(t;i|j)P_{\mathbf{x}}(j,t) \right] \tag{25}$$

which is called the *master equation* (Haken 1983; Weidlich/Haag 1983; Helbing 1995). Chiang (1968: 116ff), Lancaster (1990: pp. 109ff), and Courgeau/Lelièvre (1992: 40) presented related considerations for *continuous-time Markov processes* leading to the *Kolmogorov differential equation*

$$\frac{d}{dt}P_{\mathbf{x}}(j,t|i_0,t_0) = \sum_{i=1}^{N} \lambda_{\mathbf{x}}(t;j|i)P_{\mathbf{x}}(i,t|i_0,t_0) \quad \text{with} \quad \lambda_{\mathbf{x}}(t;j|j) := -\sum_{\substack{i=1 \\ (i \neq j)}}^{N} \lambda_{\mathbf{x}}(t;i|j). \tag{26}$$

As known from the theory of stochastic processes, the master equation can be obtained from (26) by multiplication with the initial distribution of states $P_{\mathbf{x}}(i_0,t_0)$ and subsequent summation over the initial states i_0 because of $P_{\mathbf{x}}(j,t|i_0,t_0)P_{\mathbf{x}}(i_0,t_0) = P_{\mathbf{x}}(j,t;i_0,t_0) = P_{\mathbf{x}}(i_0,t_0|j,t)P_{\mathbf{x}}(j,t)$ and (21).

According to the master equation, the temporal change of the probability $P_{\mathbf{x}}(j,t)$ to have state j at time t is given by the sum of the *effective* transition rates $\lambda_{\mathbf{x}}(t;j \leftarrow i)$ from other states i to state j minus the sum of the effective transition rates $\lambda_{\mathbf{x}}(t;i \leftarrow j)$ from state j to other states i. The *effective transition rate* $\lambda_{\mathbf{x}}(t;j \leftarrow i) := \lambda_{\mathbf{x}}(t;j|i)P_{\mathbf{x}}(i,t)$ from state i to state j is obviously the product of the transition rate $\lambda_{\mathbf{x}}(t;j|i)$ *given that* the indiviual or system under consideration is in state i times the probability $P_{\mathbf{x}}(i,t)$ that he/she/it is *actually* in state i.

In the following we will assume the special case that we have the dependence

$$\lambda_{\mathbf{x}}(t;j|i) := \lambda_{\mathbf{x}}^0(t)w_{\mathbf{x}}(j|i) \tag{27}$$

of the transition rates. Often one even restricts oneself to the *proportional hazard rate model (proportional hazards model)*

$$\lambda_{\mathbf{x}}(t;j|i) := \lambda_0(t)\exp(\boldsymbol{\beta}_{ji}\mathbf{x}) \tag{28}$$

so that $\lambda_{\mathbf{x}}^0(t) = \lambda_0(t)$ with the *baseline hazard function* $\lambda_0(t)$ and $w_{\mathbf{x}}(j|i) = \exp(\boldsymbol{\beta}_{ji}\mathbf{x})$. The optimal *parameter vectors* $\boldsymbol{\beta}_{ji}$ can be estimated from life-table data via the *partial likelihood method* proposed by Cox (1975) (see also Blossfeld et al. 1986; Kalbfleisch/Prentice 1980; Elandt-Johnson/Johnson 1980; Diekmann/Mitter 1984; Tuma/Hannan 1984; Cox/Oakes 1984; Lancaster 1990; Courgeau/Lelièvre 1992).

For reasons of simplicity we utilize relation (27) to introduce the subpopulation-specific times

$$\tau_{\mathbf{x}}(t) := \int_{t_0}^{t} dt' \, \lambda_{\mathbf{x}}^0(t') \qquad \text{corresponding to} \qquad \frac{d\tau_{\mathbf{x}}}{dt} = \lambda_{\mathbf{x}}^0(t). \tag{29}$$

This implies, for example,

$$\frac{d}{dt}P_{\mathbf{x}}(j,t) = \frac{d}{dt}P_{\mathbf{x}}[j,\tau_{\mathbf{x}}(t)] = \left(\frac{d}{d\tau_{\mathbf{x}}}P_{\mathbf{x}}(j,\tau_{\mathbf{x}})\right)\frac{d\tau_{\mathbf{x}}}{dt} = \left(\frac{d}{d\tau_{\mathbf{x}}}P_{\mathbf{x}}(j,\tau_{\mathbf{x}})\right)\lambda_{\mathbf{x}}^0(t). \tag{30}$$

As a consequence, we can bring the master equation into the form

$$\frac{d}{d\tau_{\mathbf{x}}} P_{\mathbf{x}}(j, \tau_{\mathbf{x}}) = \sum_{\substack{i=1 \\ (i \neq j)}}^{N} \left[w_{\mathbf{x}}(j|i) P_{\mathbf{x}}(i, \tau_{\mathbf{x}}) - w_{\mathbf{x}}(i|j) P_{\mathbf{x}}(j, \tau_{\mathbf{x}}) \right] \tag{31}$$

with *time-independent* transition rates

$$\lambda_{\mathbf{x}}(\tau_{\mathbf{x}}; j|i) = w_{\mathbf{x}}(j|i) . \tag{32}$$

Analogous simplifications are found for the other quantities:

$$\lambda_{\mathbf{x}}(\tau_{\mathbf{x}}|i) = w_{\mathbf{x}}(i) := \sum_{\substack{j=1 \\ (j \neq i)}}^{N} w_{\mathbf{x}}(j|i) , \tag{33}$$

$$S_{\mathbf{x}}^{n}(\tau_{\mathbf{x}}|i) = \exp\left[-w_{\mathbf{x}}(i)(\tau - \tau_{\mathbf{x}, n-1})\right] = 1 - F_{\mathbf{x}}^{n}(\tau_{\mathbf{x}}|i) = 1 - \sum_{\substack{j=1 \\ (j \neq i)}}^{N} F_{\mathbf{x}}^{n}(\tau_{\mathbf{x}}; j|i) , \tag{34}$$

and

$$f_{\mathbf{x}}^{n}(\tau_{\mathbf{x}}; j|i) = w_{\mathbf{x}}(j|i) S_{\mathbf{x}}^{n}(\tau_{\mathbf{x}}|i) = \frac{d}{d\tau_{\mathbf{x}}} F_{\mathbf{x}}^{n}(\tau_{\mathbf{x}}; j|i) , \tag{35}$$

where we have used the convention $\tau_{\mathbf{x},n} := \tau_{\mathbf{x}}(t_n)$. Since the subscript $\mathbf{x}$ for the subpopulation (cohort) is arbitrary but fixed (time-independent) we will omit it in the following which makes the mathematical formulas easier to read.

3.2   Simulations with the Master Equation

The master equation has proved to be a very powerful and quite flexible tool for the description and simulation of stochastically behaving systems which are subject to inherent or external random influences ('fluctuations'). It has got numerous applications in physics, chemistry, biology, economics, and the social sciences (Weidlich/Haag 1983; Troitzsch 1990; Weidlich 1991; Helbing 1995). The master equation can normally not be analytically solved. However, since the master equation has the form of a linear system of ordinary differential equations, it can be easily simulated by means of the usual numerical integration algorithms (Press et al. 1992: Chap. 16).

  With today's computers these simulations are very fast, so that even systems with some dozens of variables can be solved within a few minutes. Therefore, the master equation technique facilitates an efficient evaluation of all *cross-sectional quantities* related with stochastic processes. This includes the temporal course of the distribution of states $P(j, \tau)$, its maxima, means value, variance, etc. Whereas the mean value represents the average behavior of a huge number of time series, the course of the most probable time series will often (but not always) be close to the maxima of the distribution of states, and the variance is a measure for the grade of variation between different time series. In the case of a multimodal distribution of states, the mean value can considerably differ from the most probable time series.

  In summary, simulations of the master equation are suitable for *scenario techniques*, since they allow

1. to investigate and compare the properties of different conceivable stochastic models (which are normally related to different functional forms of the hazard rates),

2. to evaluate the implications of parameter changes (which may correspond to considered or planned modifications of legal regulations or social conditions),

3. to make prognoses of the range of probable future behaviors of a stochastic system.

4 Simulation of Path-Related Quantities

4.1 Microsimulations with the Monte-Carlo Technique

Although the master equation technique is very versatile, it does not allow investigations which are directly related with the *longitudinal time series*. However, the simulation of these is desirable for a number of reasons:

1. The single time series give an impression of the possible time-dependent behavior of individual systems. If the variation among different time series is small, their predictive value is large. However, if the state changes of the individual time series are large and frequent, they are normally not representative of a system's behavior. Note that, by means of the time-dependent distribution of states (i.e. by simulation of the master equation alone), we cannot always distinguish between a strongly varying (e.g. oscillatory) system behavior and a "smooth" one.

2. One can often distinguish *desirable* and *undesired* (maybe catastrophic) system states. In such cases we may be interested in the *occurence probability* with which the system under consideration up to time t takes a *desired time series* (containing desirable states only). This is relevant for *prognoses* as well as for the *controllability* of technical systems.

3. Sometimes one also likes to know the *occurence time distribution* of a certain sequencing of states.

4. When occupying an undesired state, it may be interesting to evaluate the expected *escape time* until the system reaches, for the first time, one of the desired states.

All these quantities can be obtained by evaluating a large number of stochastic time series. These can be generated and investigated in *microsimulations* applying the *Monte-Carlo simulation method* (cf. Binder 1979: Chap. 1.3.1).

The Monte-Carlo technique utilizes the fact that a system under consideration stays in an occupied state i_{n-1} for a time period $\tau_n - \tau_{n-1}$ which is exponentially distributed according to (34). Therefore, in Monte-Carlo simulations the respective life time $\tau_n - \tau_{n-1}$ is determined via the formula

$$\tau_n - \tau_{n-1} := -\frac{1}{w(i_{n-1})} \ln y_n \,, \tag{36}$$

where $y_n \in [0, 1]$ are uniformly distributed random numbers (cf. Press et al. 1992: Chap. 7). After this time period (i.e. at time τ_n) the system goes over into one of the other states

$i_n \neq i_{n-1}$ with probability

$$p(i_n|i_{n-1}) := \frac{w(i_n|i_{n-1})}{\sum\limits_{\substack{j=1 \\ (j \neq i_{n-1})}}^{N} w(j|i_{n-1})} = \lim_{\Delta\tau \to 0} \frac{P(i_n, \tau_n + \Delta\tau|i_{n-1}, \tau_n)}{\sum\limits_{\substack{j=1 \\ (j \neq i_{n-1})}}^{N} P(j, \tau_n + \Delta\tau|i_{n-1}, \tau_n)} \,. \qquad (37)$$

The realized state i_n is again selected by means of a uniformly distributed random variable $z_n \in [0,1]$: The randomly resulting value z_n corresponds to an occupation of the state i_n for which

$$Z(i_{n-1}) < z_n \leq Z(i_n) \qquad \text{with} \qquad Z(i_n) = P(j \leq i_n|i_{n-1}) := \sum_{j=1}^{i_n} p(j|i_{n-1}) \qquad (38)$$

is fulfilled. The initial state i_0 is usually given, but alternatively it can also be randomly chosen in a similar way.

Time series which are generated by means of the above outlined method have the meaning of possible realizations of the stochastic process under consideration. They can be investigated in different ways. For illustrative reasons, let us assume to have generated A different time series $Y_\alpha(\tau)$. If n_j is the number of systems found in state j at time τ, the distribution of states corresponds to

$$P(j, \tau) = \lim_{A \to \infty} \frac{n_j(\tau)}{A} \qquad \text{since} \qquad \sum_{j=1}^{N} n_j(\tau) = A \,. \qquad (39)$$

For finite A, the probability $P(n_1, \ldots, n_N; \tau)$ to find the *occupation numbers* n_j at time τ is given by the *multinomial distribution*

$$P(n_1, \ldots, n_N; \tau) = \frac{A!}{n_1! \ldots n_N!} \prod_{j=1}^{N} P(j, \tau)^{n_j} \qquad (40)$$

(cf. Helbing 1995: 70), so that the distribution of states $n_j(\tau)/A$ obtained by microsimulations will normally differ from $P(j, \tau)$. This shows that we would need a tremendous number A of simulation runs to determine $P(j, \tau)$ from microsimulations. For other quantities the situation is similar. As a consequence, the Monte-Carlo technique is rather a brute force than an efficient method. It is, however, very suitable for generating some sample time series for illustrative purposes (which is sufficient for Item 1). The big advantage of the microsimulation technique is that it usually allows a relatively simple treatment of systems with a huge number of possible system states and/or very complex relations for the transition rates. Therefore, it is mainly used in situations where the derivation or simulation of a master equation is too difficult.

Since the microsimulation technique normally requires an extreme computational effort, it should only be applied if there are no suitable alternatives. As discussed in Section 3.2, the distribution of states $P(j, \tau)$ is much easier obtained by simulating the master equation. In the next section it will be shown that more efficient methods than the microsimulation technique can also be developed for the problems raised in Items 2, 3, and 4. The reason is, that they only regard the *sequencing* of time series up to a certain time τ but not the *waiting times* after which the single state of the time series are occupied.

4.2 Occurence Probabilities and Occurence Times of Paths

With the results of sections 2 and 3 we can now calculate the probability $P(H_n, t)$ that we have the event history

$$H_n := \{t_0, i_0; t_1, i_1; \ldots; t_n, i_n\} = \{\tau_0, i_0; \tau_1, i_1; \ldots; \tau_n, i_n\} \tag{41}$$

at time $\tau(t)$. Remembering that $S^k(\tau_k|i_{k-1}) = P(\mathcal{T}_k \geq \tau_k|Y_{k-1} = i_{k-1})$ with

$$\mathcal{T}_k := \tau(T_k) \tag{42}$$

is the probability to stay in state i_{k-1} up to time τ_k and that $w(i_k|i_{k-1})d\tau_k = P(i_k, \tau_k + d\tau_k|i_{k-1}, \tau_k)$ is the probability to change from state i_{k-1} to state i_k between times τ_k and $\tau_k + d\tau_k$, we find

$$
\begin{aligned}
P(H_n, \tau) & = S^{n+1}(\tau|i_n)w(i_n|i_{n-1})d\tau_n\, S^n(\tau_n|i_{n-1})w(i_{n-1}|i_{n-2})d\tau_{n-1} \\
& \quad \ldots w(i_2|i_1)d\tau_2\, S^2(\tau_2|i_1)w(i_1|i_0)d\tau_1\, S^1(\tau_1|i_0)P(i_0, \tau_0)
\end{aligned} \tag{43}
$$

(Empacher 1992). Often, however, one is not interested in the times $\tau_k = \tau(t_k)$ at which the single transitions occur, but only in the sequencing, i.e. in the *path*

$$\mathcal{C}_n := i_n \longleftarrow i_{n-1} \longleftarrow \ldots \longleftarrow i_1 \longleftarrow i_0 \tag{44}$$

which the system has taken up to time $\tau(t)$. For example, one could ask which is the probability that somebody is married the second time or unemployed the third time at time τ. One could also compare the probabilities of being married the second time after divorce or after death of the partner. The necessary quantities for answering questions like these can be derived from formula (43) by integration with respect to $\tau_1, \ldots, \tau_n$. For the probability of having path $\mathcal{C}_n$ at time τ we obtain

$$P(\mathcal{C}_n, \tau) = \sum_{k=0}^{n} \frac{e^{-w_k\tau}}{\displaystyle\prod_{\substack{l=0 \\ (w_l \neq w_k)}}^{n} (w_l - w_k)} \, p_{m_k}(w_k, \tau)w(\mathcal{C}_n)P(i_0, t_0)\,. \tag{45}$$

Here, m_k is the *multiplicity* of the overall departure rate $w_k := w(i_k)$ in path $\mathcal{C}_n$,

$$w(\mathcal{C}_n) = w(i_n \leftarrow i_{n-1} \leftarrow \ldots \leftarrow i_0) := \begin{cases} 1 & \text{if } n = 0 \\ \displaystyle\prod_{l=1}^{n} w(i_l|i_{l-1}) & \text{if } n \geq 1, \end{cases} \tag{46}$$

$$
\begin{aligned}
p_m & = \frac{(-1)^{m+1}}{m(m-1)}\Bigg(g^{(m-1)} + \sum_{n_1=1}^{m-2} \frac{g^{(m-1-n_1)}}{n_1}\bigg(g^{(n_1)} + \sum_{n_2=1}^{n_1-1} \frac{g^{(n_1-n_2)}}{n_2} \\
& \quad \times \bigg(g^{(n_2)} + \ldots + \sum_{n_{n_0-1}=1}^{n_{n_0-2}-1} \bigg(\frac{g^{(n_{n_0-2}-n_{n_0-1})}}{n_{n_0-1}} g^{(n_{n_0-1})}\bigg)\ldots\bigg)\bigg)\Bigg)
\end{aligned} \tag{47}
$$

for $m \geq 2$, $p_m = 1$ for $m = 1$, and

$$g^{(l+1)}(w_k, \tau) := \sum_{\substack{l=0 \\ (w_l \neq w_k)}}^{n} \frac{1}{(w_l - w_k)^{l+1}} - \tau\delta_{l0} \tag{48}$$

with $\delta_{l0} = 1$ if $l = 0$ and $\delta_{l0} = 0$ otherwise. The detailed steps on the way of deriving $P(\mathcal{C}_n, \tau)$ are presented in a paper by Helbing and Molini (1995). If one restricts to the case that all overall departure rates w_k are different from each other (i.e. to pure *birth processes*) the above relation simplifies to

$$
P(\mathcal{C}_n, \tau) = \sum_{k=0}^{n} \frac{S^1(\tau|i_k)}{\prod_{\substack{l=0 \\ (l \neq k)}}^{n} [w(i_l) - w(i_k)]} w(\mathcal{C}_n) P(i_0, \tau_0) \,, \quad \text{where} \quad S^1(\tau|i_k) = \mathrm{e}^{-w(i_k)(\tau - \tau_0)} \,.
$$

(49)

The special formula (49) was already presented by Chiang (1968: 50ff) who, however, did not introduce the more general path concept developed above.

The *probability density $P(\tau|\mathcal{C}_n)$ of the occurence times τ* (the *occurence time distribution*) of path $\mathcal{C}_n$ can now easily be obtained from (45). It is given by the formula

$$
P(\tau|\mathcal{C}_n) = \frac{P(\mathcal{C}_n, \tau)}{P(\mathcal{C}_n)} \quad \text{with} \quad P(\mathcal{C}_n) := \int_{\tau_0}^{\infty} d\tau \, P(\mathcal{C}_n, \tau) = w(\mathcal{C}_n) P(i_0, \tau_0) \prod_{k=0}^{n} \frac{1}{w(i_k)}
$$

(50)

(Helbing, 1994, 1995). From this we can derive the *average*

$$
\langle \tau \rangle_{\mathcal{C}_n} = \int_{\tau_0}^{\infty} d\tau \, \tau P(\tau|\mathcal{C}_n) = \sum_{k=0}^{n} \frac{1}{w(i_k)}
$$

(51)

of the occurence times τ and their *variance*

$$
\Theta_{\mathcal{C}_n} = \left\langle (\tau - \langle \tau \rangle_{\mathcal{C}_n})^2 \right\rangle_{\mathcal{C}_n} = \sum_{k=0}^{n} \frac{1}{[w(i_k)]^2}
$$

(52)

(cf. Helbing 1994, 1995). The average occurence times are the basis for the calculation of the expected *escape times.*

4.3   The Simulation Tool EPIS

For a determination of the quantities called for by Items 2 and 4 of Section 4.1 we have to sum up over the occurence probabilities or average occurence times of many paths. In the following, we will discuss how this can be done in an efficient way with respect to computer memory and simulation time. For illustrative reasons we begin with the socalled *contracted path integral solution* of the master equation. This reads

$$
P(j, \tau) = \sum_{n=0}^{\infty} \sum_{\mathcal{C}_n} P(\mathcal{C}_n, \tau)
$$

(53)

with

$$
\sum_{\mathcal{C}_n} := \sum_{\substack{i_{n-1}=1 \\ (i_{n-1} \neq j)}}^{N} \sum_{\substack{i_{n-2}=1 \\ (i_{n-2} \neq i_{n-1})}}^{N} \cdots \sum_{\substack{i_0=1 \\ (i_0 \neq i_1)}}^{N} \,,
$$

(54)

since, using $\tau_0 = 0$, the probability $P(j, \tau)$ to find state j at time τ is given as the sum over the occurence probabilities $P(\mathcal{C}_n, \tau)$ of all paths $\mathcal{C}_n$ with an arbitrary length n which lead to state $i_n := j$ within the time interval τ (Helbing 1994, 1995).

When (53) is evaluated numerically, one must restrict the summation to a finite number of *relevant* paths. Here, we can utilize the fact that the occurence probability $P(\mathcal{C}_n, \tau)$ of a path $\mathcal{C}_n$ with a non-absorbing final state j is negligible if

$$|\tau - \langle\tau\rangle_{\mathcal{C}_n}| \le a\sqrt{\Theta_{\mathcal{C}_n}} \tag{55}$$

with a suitably chosen *accuracy parameter* a. For $\tau < \langle\tau\rangle_{\mathcal{C}_n} - a\sqrt{\Theta_{\mathcal{C}_n}}$ there is not enough time to traverse all states i_k of the path, whereas for $\tau > \langle\tau\rangle_{\mathcal{C}_n} + a\sqrt{\Theta_{\mathcal{C}_n}}$ the system will probably have already visited additional states i_{n+1}, i_{n+2}, ... Therefore, only paths which fulfil condition (55) are *relevant*. Choosing the accuracy parameter $a \approx 3$ allows to reconstruct about 99% of the probability distribution $P(j, \tau)$ which can be checked by means of the *normalization condition* $\sum_{j=1}^{N} P(j, \tau) = 1$.

If we are interested in the probability that the system takes *desired* paths only, we have to take into account an even smaller number of relevant paths. This is simply done by restricting the summation (54) to the desired states. A similar procedure is used to determine the expected escape time of a system from undesired states.

In order to evaluate the path-related quantities introduced above, the simulation tool EPIS (**E**fficient **P**ath **I**ntegral **S**imulator) has recently been developed at the University of Stuttgart (Molini 1995). This bases on the standard path-search algorithm *depth-first* (Schildt 1990) but restricts it to the relevant paths (see Figure 2). The advantage of this algorithm is that it uses computer memory very efficiently and allows to calculate certain quantities like

$$\langle\tau\rangle_{\mathcal{C}_{n+1}} = \langle\tau\rangle_{\mathcal{C}_n} + \frac{1}{w_{n+1}}, \quad \Theta_{\mathcal{C}_{n+1}} = \Theta_{\mathcal{C}_n} + \frac{1}{(w_{n+1})^2}, \quad w(\mathcal{C}_{n+1}) = w(\mathcal{C}_n)w(i_{n+1}|i_n) \quad (56)$$

after each step. In this way, the results of previous calculations can be utilized which again saves computer time.

The simulation tool EPIS is constructed for systematically generating relevant paths and for calculating

- their occurence probabilities (45),

- their occurence time distributions (50),

- the occurence probability of desired paths up to time τ,

- the expected escape time from undesired states.

Moreover, the contracted path integral solution (53) can be compared with the distribution $P(j, \tau)$ of states which results by a numerical integration of the master equation.

5 Effective Cumulative Life-Time Distribution and Hidden State Concept

The time-dependence of hazard rates is usually fitted to functions like the ones belonging to the Gompertz distribution, the Weibull distribution, the log-normal distribution, the log-logistic distribution, the sickle distribution, the extreme-value distribution,
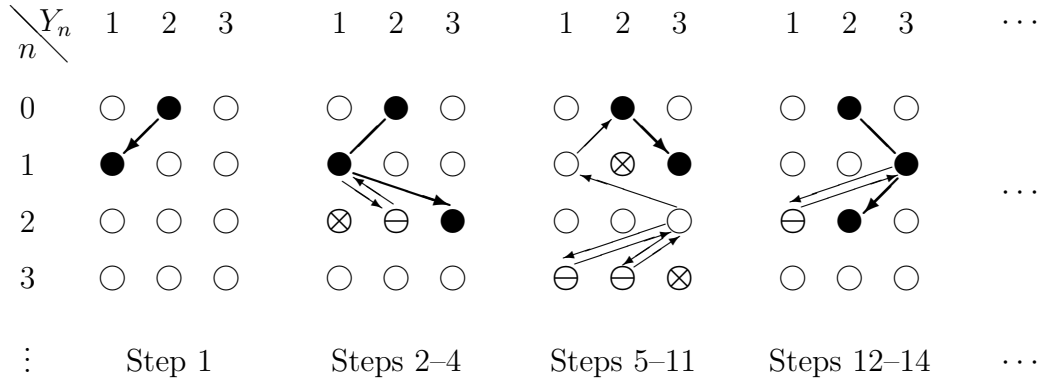
Figure 2: Efficient path generation according to the modified *depth-first* algorithm. Here, it is assumed that the system can be in one of the states $Y_n \in \mathcal{M} = \{1, 2, 3\}$ after the nth transition and that it is initially in the state $Y_0 = i_0 = 2$. The single steps of the procedure are symbolized by arrows. They correspond to an extension of a path $\mathcal{C}_n$ by a new admissible state $Y_{n+1} = i_{n+1}$ or to a removal of the last state i_n if the path $\mathcal{C}_n$ cannot be further extended to a relevant path $\mathcal{C}_{n+1}$. Full circles represent the states of the last generated path and are connected by thick arrows or lines. Crossed circles ($\otimes$) stand for states Y_{n+1} which are not admissible since transitions from state Y_n must lead to another state $Y_{n+1} \neq Y_n$. The symbol "$\ominus$" indicates that the resulting path $\mathcal{C}_{n+1}$ turns out to be not relevant due to $\tau > \langle \tau \rangle_{\mathcal{C}_{n+1}} + a\sqrt{\Theta_{\mathcal{C}_{n+1}}}$. Since then all longer paths are irrelevant, too, the procedure removes the last state and tries to extend the remaining path $\mathcal{C}_n$ by another state. If this is not possible, the state i_n is also removed, etc.

the gamma distribution, and others (Blossfeld et al. 1986; Kalbfleisch/Prentice 1980; Elandt-Johnson/Johnson 1980; Tuma/Hannan 1984; Cox/Oakes 1984; Lancaster 1990; Courgeau/Lelièvre 1992). Since these functional time-dependences can not always be interpreted in terms of underlying social or psychological processes, a new concept for an *interpretation* of the respective time-dependence of transition rates is proposed in the following. This bases on the discovery that, if we would have sequential or parallel transitions between *hidden states*, this would cause a time-dependence of the related *effective* transition rate $w_{\mathrm{eff}}(\tau; j|i)$ of the *total* transition process. Hidden states denote states which are either not phenomenologically distinguishable from other states, not externally measurable, or simply not detected. They could, for example, reflect the psychological and mental stages preceding a concrete action, the individual predisposition, or the internal attitude towards a certain behavior.

It should be underlined that the aim of the hidden state concept is not to question or criticize but to supplement the powerful methods of classical survival analysis, which have proved their suitability in numerous applications. Favourable properties of the hidden state concept are

1. the huge number of classes of effective transition rates which can be generated by it and which may include additional kinds of time-dependences,

2. the theoretical relationship between these different classes of time-dependent transition rates,

3. the possibility of a direct interpretation of each hidden state model.

13

Before the hidden state concept is discussed in general (cf. Section 5.3), it will be illustrated by two well-known cases which correspond to *pure birth processes* (cf. Section 5.1) and to the situation of *unobserved heterogeneity* (cf. Section 5.2), respectively. This allows to introduce the basic ideas without complex mathematical considerations which are unavoidable in the general case.

5.1 Sequential Transitions

As an example for a series of birth processes, let us face the transition

$$\text{married} \longrightarrow \text{divorced} \tag{57}$$

and assume that it is actually a transition of the kind

$$\text{married} \longrightarrow \text{married with attitude towards being divorced} \longrightarrow \text{divorced}\,. \tag{58}$$

It should be mentioned that this example is, of course, only a "toy model" which was chosen exclusively for illustrative reasons.

In order to take into account the hidden state "married with attitude towards being divorced" we can introduce the following notation of states:

$$
\begin{aligned}
(1,1) &:= \text{married}\,,\\
(1,2) &:= \text{married with attitude towards being divorced}\,,\\
(2,1) &:= \text{divorced}\,.
\end{aligned}
\tag{59}
$$

With this, example (58) can be written in the form

$$\mathcal{H}_2 := (2,1) \longleftarrow (1,2) \longleftarrow (1,1) \tag{60}$$

or

$$\mathcal{H}_2 := \mathbf{j} \longleftarrow \mathbf{i}_1 \longleftarrow \mathbf{i}_0\,, \tag{61}$$

where $\mathbf{j} = \mathbf{i}_2 := (2,1)$, $\mathbf{i}_1 := (1,2)$, and $\mathbf{i}_0 := (1,1)$. In general, the distinguished states $i \in \{1,\ldots,N\}$ actually comprise the *"hidden states"*

$$\mathbf{i} := (i,h) \qquad \text{with} \qquad h \in \{1,\ldots,N_i\}\,, \tag{62}$$

where h differentiates the (possibly unknown) variants of state i.

Analogous to (12) the *effective cumulative life-time distribution* $F_{\text{eff}}^1(\tau|i)$ for the *total sequence* (61) of transitions can be defined by the probability $P(\mathcal{T}_2 < \tau|\mathbf{i}_1 \leftarrow \mathbf{i}_0)$ that the final transition (i.e. the transition into state $\mathbf{j}$) takes place before time τ given that the preceding path was $\mathbf{i}_1 \longleftarrow \mathbf{i}_0$. Since $P(\mathcal{H}_2,\tau)$ is the probability to have path $\mathcal{H}_2$ at time τ we get

$$F_{\text{eff}}^1(\tau|i) = P(\mathcal{T}_2 < \tau|\mathbf{i}_1 \leftarrow \mathbf{i}_0) = P(\mathcal{H}_2,\tau)\,. \tag{63}$$

Inserting (49) leads to

$$
\begin{aligned}
F_{\text{eff}}^1(\tau|i) = \Bigg(&\frac{S^1(\tau|\mathbf{i}_0)}{-w(\mathbf{i}_0)[w(\mathbf{i}_1) - w(\mathbf{i}_0)]} + \frac{S^1(\tau|\mathbf{i}_1)}{-w(\mathbf{i}_1)[w(\mathbf{i}_0) - w(\mathbf{i}_1)]}\\
&+ \frac{1}{w(\mathbf{i}_1)w(\mathbf{i}_0)}\Bigg)\, w(\mathbf{i}_1)w(\mathbf{i}_0)
\end{aligned}
\tag{64}
$$

because of $w(\mathbf{j}) = 0$, $w(\mathbf{j}|\mathbf{i}_1) = w(\mathbf{i}_1)$, $w(\mathbf{i}_1|\mathbf{i}_0) = w(\mathbf{i}_0)$, and $P(\mathbf{i}_0, \tau_0) = 1$. With this we can define the *effective survivor function*

$$S^1_{\text{eff}}(\tau|i) := 1 - F^1_{\text{eff}}(\tau|i) = \left(\frac{S^1(\tau|\mathbf{i}_0)}{w(\mathbf{i}_0)} - \frac{S^1(\tau|\mathbf{i}_1)}{w(\mathbf{i}_1)} \right) \frac{w(\mathbf{i}_1)w(\mathbf{i}_0)}{w(\mathbf{i}_1) - w(\mathbf{i}_0)} \tag{65}$$

in accordance with (34).

In the following we try to represent this expression in the form

$$S^1_{\text{eff}}(\tau|i) = \exp\left[-\int_{\tau_0}^{\tau} d\tau' \, w_{\text{eff}}(\tau'|i) \right] \tag{66}$$

(cf. (19)) which corresponds to the *detected (observed, measured)* transition (57) of the type $j \longleftarrow i$ (where i corresponds to "married" and j to "divorced"). The appropriate relation for the *effective overall departure rate* $w_{\text{eff}}(\tau|i)$ is found by differentiation of (66) with respect to τ. We obtain

$$\frac{d}{d\tau} S^1_{\text{eff}}(\tau|i) = -w_{\text{eff}}(\tau|i) S^1_{\text{eff}}(\tau|i) \qquad \text{i.e.} \qquad w_{\text{eff}}(\tau|i) = -\frac{1}{S^1_{\text{eff}}(\tau|i)} \frac{d}{d\tau} S^1_{\text{eff}}(\tau|i). \tag{67}$$

Inserting (65) finally gives

$$w_{\text{eff}}(\tau|i) = \frac{S^1(\tau|\mathbf{i}_0) - S^1(\tau|\mathbf{i}_1)}{\dfrac{S^1(\tau|\mathbf{i}_0)}{w(\mathbf{i}_0)} - \dfrac{S^1(\tau|\mathbf{i}_1)}{w(\mathbf{i}_1)}}. \tag{68}$$

For the *effective transition rate* $w_{\text{eff}}(\tau; j|i)$ we find $w_{\text{eff}}(\tau; j|i) = w_{\text{eff}}(\tau|i)$ since there are no alternative transitions from state i to state j.

The most important fact about the relations for $w_{\text{eff}}(\tau; j|i)$ and $w_{\text{eff}}(\tau|i)$ is that these are *time-dependent.* Therefore, the hidden state concept of behavioral changes may serve as a means for interpreting the time-dependence of hazard rates. A comparison with empirical data, however, shows that the above model (58) is still too simple for an explanation of the sickle curve which is found for the hazard rate of divorces (Diekmann/Mitter 1984).

The above toy model could be improved by distinguishing different motivations (i.e. alternative reasons) which may lead to a divorce. This brings us to hidden state models with parallel transitions which correspond to cases of unobserved heterogeneity.

5.2   Parallel Transitions

For illustrative reasons we will discuss an example which stems from Kinsey et al. (1948) concerning the kinds of sexual activities of white males in the United States. It should be mentioned that this rather old example is not intended to demonstrate survival analytical or statistical methods. In addition, since it is based on a small amount of cross-sectional data, it only allows the determination of the most frequent transitions. Longitudinal data would, of course, supply more detailed information, e.g. concerning the possible (but seemingly infrequent) transitions "heterosexual activity" $\longrightarrow$ "bisexual activity" or "bisexual activity" $\longrightarrow$ "homosexual activity". Therefore, the interpretation of the data will also not be the focus of our discussion. The example was rather chosen for didactical reasons, since it is suitable for illuminating

1. the hidden state model with exclusively parallel transitions,

2. possible methods for the detection and separation of hidden variables (here: two different kinds of "homosexual activity") from empirical data,

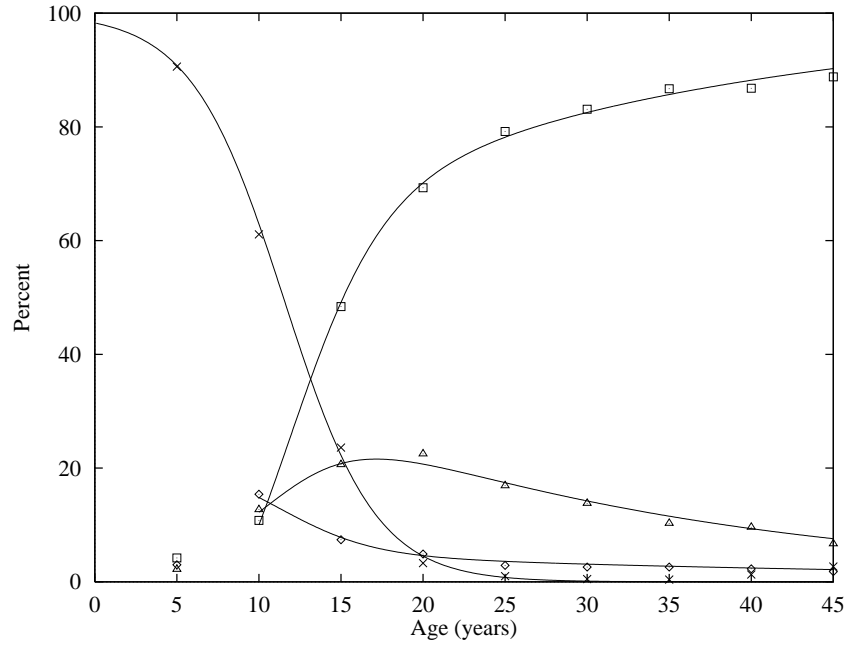3. a simulation model basing on the master equation.

Figure 3: Percent of white males engaging in heterosexual activity ($\square$), bisexual activity ($\triangle$), homosexual activity ($\diamond$), or having no socio-sexual contacts ($\times$) in dependence of age. Here, "bisexual activity" summarizes the five classes "predominantly heterosexual, only incidentally homosexual", "predominantly heterosexual, more than incidentally homosexual", "equally heterosexual and homosexual", "predominantly homosexual, more than incidentally heterosexual", and "predominantly homosexual, only incidentally heterosexual". The solid lines show the simulation results of the hidden state model which is proposed later on. Despite its simplicity it apparently fits the data quite well.

We now come to the example by Kinsey et al. (1948). Figure 3 shows the distribution of the different kinds of sexual activities (within a period of three years) in dependence of age τ. It turns out that the proportion

$$X_{\text{tot}}(\tau) := \sum_{j=1}^{3} X_j(\tau) \tag{69}$$

of males with socio-sexual contacts can be approximated by the *logistic equation* (Pearl 1924; Verhulst 1845)

$$\frac{dX_{\text{tot}}}{d\tau} = \nu X_{\text{tot}}(\tau)[1 - X_{\text{tot}}(\tau)], \tag{70}$$

where $X_{\text{tot}}(0 \text{ years}) = 0.0173$, $\nu = 0.36$ per year, and

$X_1(\tau) \quad := \quad$ proportion of white males engaging in heterosexual activity $(j = 1)$,

$$X_2(\tau) \quad := \quad \text{proportion of white males engaging in bisexual activity } (j = 2)\,,$$
$$X_3(\tau) \quad := \quad \text{proportion of white males engaging in homosexual activity } (j = 3)\,. \quad (71)$$

Equation (70) could be interpreted in the way that the proportion $1 - X_{\text{tot}}$ of males *without* socio-sexual contacts is seduced to sexual contacts by the proportion X_{tot} of individuals of about the same age *with* socio-sexual contacts.

However, note that the logistic curve deviates from the empirical data for very young and for old males. Whereas the proportion of males without socio-sexual contacts should start with a value of 1 at the age of 0 years, it should increase after an age of about 30 years due to the decrease of sexual opportunities (or other reasons). In any case, empirical and theoretical statements are very questionable before an age of 10 or 15 years.

We will now investigate the proportions

$$P(j, \tau) := \frac{X_j(\tau)}{X_{\text{tot}}(\tau)} \quad (72)$$

of white males *with* socio-sexual contacts who engage in sexual activity j. These are represented half-logarithmically in Figure 4. $P(j, \tau)$ can also be interpreted as probability that, in a representative sample of males with socio-sexual contacts, a randomly picked out male engages in sexual activity j.
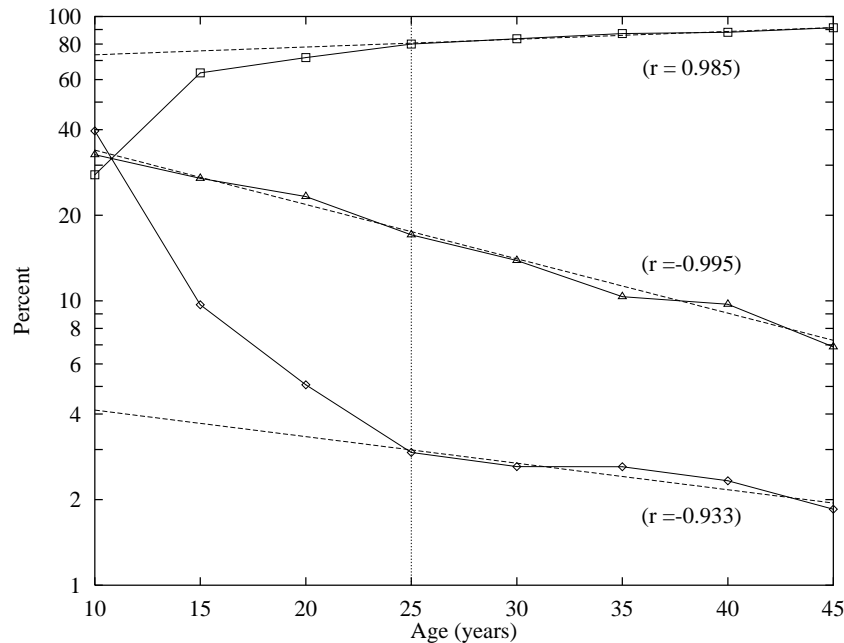

Figure 4: Half-logarithmic plot of the proportions of sexually active males engaging in heterosexual activity ($\square$), bisexual activity ($\triangle$), and homosexual activity ($\diamond$). The broken lines show the corresponding linear regression curves (on the basis of the time period 10 years $\leq \tau \leq 45$ years for bisexual activity and the time period 25 years $\leq \tau \leq 45$ years otherwise). The numbers in brackets are the respective correlation coefficients.

We find that heterosexual activity increases, but bisexual and homosexual activity decrease with advancing age which might be a consequence of an adaptation to social norms. The half-logarithmically scaled curve of bisexual activity is quite well decribed by

a linear relation over the whole range of ages, whereas the other curves behave almost linearly after an age of 25 years. Between 10 and 25 years there seems to be a surplus of homosexual activity connected with less heterosexual activity than expected. This surplus decays very fast and fills the gap of heterosexual activity. We may therefore suspect to be confronted with two different types of homosexual activity:

1. a rather permanent type of homosexual activity (with a half-life period of about 30 years) which might arise from a "homosexual predisposition",

2. a short-lived type of "adolescent homosexual activity" (with a half-life period of about 2 years) which might be a substitute for lacking opportunities for heterosexual activity.

This interpretation is in good agreement with the findings of developmental psychology.

Summarizing the previous considerations we predominantly have the following transitions

$$
\begin{array}{l}
(2,1) \searrow \\
(3,1) \longrightarrow (1,1) \\
(3,2) \nearrow
\end{array}
$$
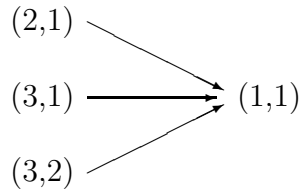
between the states

$$
\begin{aligned}
\mathbf{j} := (1,1) &:= \text{"heterosexual activity"} , \\
\mathbf{i}_0^1 := (2,1) &:= \text{"bisexual activity"} , \\
\mathbf{i}_0^2 := (3,1) &:= \text{"homosexual activity with homosexual predisposition"} , \\
\mathbf{i}_0^3 := (3,2) &:= \text{"adolescent homosexual activity, but heterosexual predisposition"} ,
\end{aligned}
\tag{73}
$$

where we have again used the hidden state representation (62). As Figure 3 shows, the temporal course of the proportions $X_j(\tau)$ can, for $\tau \geq \tau_0 := 10$ years, be approximated by (70) and (72) together with the master equation

$$
\frac{d}{dt}P(\mathbf{i}',t) = \sum_{\mathbf{i}(\neq \mathbf{i}')} \left[w(\mathbf{i}'|\mathbf{i})P(\mathbf{i},t) - w(\mathbf{i}|\mathbf{i}')P(\mathbf{i}',t) \right] \qquad (\mathbf{i},\mathbf{i}' \in \{\mathbf{j},\mathbf{i}_0^1,\mathbf{i}_0^2,\mathbf{i}_0^3\}) . \tag{74}
$$

The presented simulation results are for the parameter values

$$
\begin{aligned}
P(\mathbf{i}_0^1,\tau_0) &= 0.33 , & w(\mathbf{j}|\mathbf{i}_0^1) &= 0.042 \text{ per year} , \\
P(\mathbf{i}_0^2,\tau_0) &= 0.05 , & w(\mathbf{j}|\mathbf{i}_0^2) &= 0.024 \text{ per year} , \\
P(\mathbf{i}_0^3,\tau_0) &= 0.35 , & w(\mathbf{j}|\mathbf{i}_0^3) &= 0.366 \text{ per year}
\end{aligned}
\tag{75}
$$

which were determined with the *method of least squares* (Elandt-Johnson/Johnson 1980; Tuma/Hannan 1984; Lancaster 1990; Helbing 1995). Although the terms $w(\mathbf{i}'|\mathbf{i})P(\mathbf{i},t)$ were assumed to be negligible for $\mathbf{i}' \neq \mathbf{j}$ in accordance with the simplified model under consideration, the simulation results fit the empirical data quite well.

A number of alternative models with the same number of parameters have also been tested. Most of them yielded worse correlations with the empirical data. Only the assumption that "homosexual activity with homosexual predisposition" is replaced by "bisexual

activity" instead of "heterosexual activity" produced comparable results. A better correlation with the empirical data can, of course, always be reached by more complicated models which include a greater number of parameters. However, an increase of the number of parameters is only advisable, if this gives a considerably better fit. Otherwise, some of the parameters will be insignificant so that the explanatory power of at least some variables is very low.

In the following we want to describe the transitions from the two "hidden" homosexual states $\mathbf{i}_0^2 = (3, 1)$ and $\mathbf{i}_0^3 = (3, 2)$ in an overall manner as if we would only have one homosexual state $i = 3$. The associated effective transition scheme (which corresponds to the states that were actually detected) is

$$
\begin{array}{c}
2 \\
 \\
3
\end{array}
\searrow
\quad 1
$$
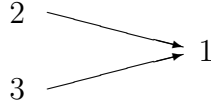
The *effective cumulative life-time distribution* $F_{\text{eff}}^1(\tau|i) = P(\mathcal{T}_1 < \tau|X_0 = i)$ for the transition from homosexual to heterosexual activity is given by the probability that one of the paths $\mathbf{j} \longleftarrow \mathbf{i}_0^2$ or $\mathbf{j} \longleftarrow \mathbf{i}_0^3$ is taken before time τ:

$$
F_{\text{eff}}^1(\tau|i) = P(\mathbf{j} \leftarrow \mathbf{i}_0^2, \tau) + P(\mathbf{j} \leftarrow \mathbf{i}_0^3, \tau). \tag{76}
$$

From (49) we obtain

$$
\begin{aligned}
F_{\text{eff}}^1(\tau|i) &= \left(\frac{S^1(\tau|\mathbf{i}_0^2)}{w(\mathbf{j}) - w(\mathbf{i}_0^2)} + \frac{S^1(\tau|\mathbf{j})}{w(\mathbf{i}_0^2) - w(\mathbf{j})} \right) w(\mathbf{j}|\mathbf{i}_0^2) P(\mathbf{i}_0^2, \tau_0) \\
&+ \left(\frac{S^1(\tau|\mathbf{i}_0^3)}{w(\mathbf{j}) - w(\mathbf{i}_0^3)} + \frac{S^1(\tau|\mathbf{j})}{w(\mathbf{i}_0^3) - w(\mathbf{j})} \right) w(\mathbf{j}|\mathbf{i}_0^3) P(\mathbf{i}_0^3, \tau_0) \\
&= [1 - S^1(\tau|\mathbf{i}_0^2)] P(\mathbf{i}_0^2, \tau_0) + [1 - S^1(\tau|\mathbf{i}_0^3)] P(\mathbf{i}_0^3, \tau_0) \tag{77}
\end{aligned}
$$

due to $w(\mathbf{j}|\mathbf{i}_0^2) = w(\mathbf{i}_0^2)$, $w(\mathbf{j}|\mathbf{i}_0^3) = w(\mathbf{i}_0^3)$, and $w(\mathbf{j}) = 0$. With $S_{\text{eff}}^1(\tau|i) = 1 - F_{\text{eff}}^1(\tau|i)$ we now find the *effective survivor function*

$$
S_{\text{eff}}^1(\tau|i) = S^1(\tau|\mathbf{i}_0^2) P(\mathbf{i}_0^2, \tau_0) + S^1(\tau|\mathbf{i}_0^3) P(\mathbf{i}_0^3, \tau_0). \tag{78}
$$

If we again try to express the effective survivor function in form (66), we can calculate the effective overall departure rate $w_{\text{eff}}(\tau|i)$ via formula (67). Inserting (78) provides

$$
w_{\text{eff}}(\tau|i) = \frac{w(\mathbf{i}_0^2) S^1(\tau|\mathbf{i}_0^2) P(\mathbf{i}_0^2, \tau_0) + w(\mathbf{i}_0^3) S^1(\tau|\mathbf{i}_0^3) P(\mathbf{i}_0^3, \tau_0)}{S^1(\tau|\mathbf{i}_0^2) P(\mathbf{i}_0^2, \tau_0) + S^1(\tau|\mathbf{i}_0^3) P(\mathbf{i}_0^3, \tau_0)}. \tag{79}
$$

This result obviously differs from expression (68) which we obtained for the hidden state model (61) of the Section 5.1. However, since we have only transitions to *one* final state, we again find $w_{\text{eff}}(\tau; j|i) = w_{\text{eff}}(\tau|i)$ for the *effective transition rates* $w_{\text{eff}}(\tau; j|i)$.

Note that the two hidden homosexual states are an example for the widespread case of *unobserved heterogeneity (unmeasured heterogeneity)* (Blossfeld et al. 1986; Elandt-Johnson/Johnson 1980; Diekmann/Mitter 1984; Tuma/Hannan 1984; Cox/Oakes 1984; Lancaster 1990; Courgeau/Lelièvre 1992). In the general case of *heterogeneous populations (mixed populations)*, we have the relations

$$
S_{\text{eff}}^1(\tau|i) = \sum_k S^1(\tau|\mathbf{i}_0^k) P(\mathbf{i}_0^k, \tau_0) \tag{80}
$$

and

$$w_{\text{eff}}(\tau; j|i) = w_{\text{eff}}(\tau|i) = \frac{\sum_k f^1(\tau; \mathbf{j}|\mathbf{i}_0^k) P(\mathbf{i}_0^k, \tau_0)}{\sum_k S^1(\tau|\mathbf{i}_0^k) P(\mathbf{i}_0^k, \tau_0)} \tag{81}$$

if there is only one final state $\mathbf{j} = (j, 1)$ (Courgeau/Lelièvre 1992: 44; Blossfeld et al. 1986: 93). Formula (80) describes the mixture of survival functions in heterogeneous populations. $P(\mathbf{i}_0^k, \tau_0)$ is called the *mixing distribution (compounding distribution)*, and the state $i = i_0$ summarizing the states $\mathbf{i}_0^k$ can be interpreted as *compound state*. In the special case that some unobserved subpopulations (which are here reflected by k) are subject to transitions but others not, we are confronted with a so-called *mover-stayer model* (cf. Courgeau/Lelièvre 1992). The reason is that the formerly mentioned subpopulations can be interpreted as *movers*, the latter ones (which are characterized by $w(\mathbf{i}_0^k) = 0$) as *stayers*.

5.3 The General Case

An investigation of formula (68) shows that *sequential* transitions are related with an effective transition rate $w_{\text{eff}}(\tau; j|i)$ which is zero at $\tau = \tau_0$ and *increases* monotonically in the course of time up to the limiting value $\min[w(\mathbf{i}_0), w(\mathbf{i}_1)]$. In contrast, for the case (79) of *parallel* transitions we find that the corresponding effective transition rate $w_{\text{eff}}(\tau; j|i)$ starts with the finite value $w_{\text{eff}}(\tau_0; j|i) = w(\mathbf{i}_0^2) P(\mathbf{i}_0^2, \tau_0) + w(\mathbf{i}_0^3) P(\mathbf{i}_0^3, \tau_0)$ and *decreases* monotonically up to the lower limit $\min[w(\mathbf{i}_0^2), w(\mathbf{i}_0^3)]$.

Therefore one can conjecture that *any* time-dependence of hazard rates (including sickle-shaped ones) can be obtained or at least approximated by a suitable combination of sequential and parallel transitions. The general hidden state concept being necessary for this is presented in the following and can be skipped by readers who are not interested in the mathematical details.

Again we assume that the detected states i comprise a number of hidden states $\mathbf{i} = (i, h)$ with $h \in \{1, 2, \dots, N_i\}$ and $N_i \geq 1$. In the general case, a detected transition $j \longleftarrow i$ is composed of all kinds of transitions between the hidden states $\mathbf{i}_k = (i, h_k)$ comprised by state i ($h_k \in \{1, \dots, N_i\}$) before the individual or system comes the first time into one of the hidden states $\mathbf{j} = (j, h_n)$ ($h_n \in \{1, \dots, N_j\}$) which are comprised by state j (see Figure 5). Therefore, we will have to sum up over all paths $\mathcal{H}$ of the form

$$\mathcal{H}_n := \mathbf{j} \longleftarrow \mathbf{i}_{n-1} \longleftarrow \dots \longleftarrow \mathbf{i}_0 \qquad \text{with} \quad \mathbf{i}_k = (i, h_k) \quad \text{and} \quad \mathbf{j} = (j, h_n), \tag{82}$$

where the path length n is arbitrary. For this purpose we introduce the abbreviation

$$\sum_{\mathcal{H}} := \sum_{n=0}^{\infty} \sum_{h_n=1}^{N_j} \sum_{\mathcal{H}_n} \qquad \text{with} \qquad \sum_{\mathcal{H}_n} := \sum_{h_{n-1}=1}^{N_i} \sum_{\substack{h_{n-2}=1 \\ (h_{n-2} \neq h_{n-1})}}^{N_i} \dots \sum_{\substack{h_0=1 \\ (h_0 \neq h_1)}}^{N_i} \tag{83}$$

similar to (54).

We are now looking for an expression which allows to calculate the probability $F_{\text{eff}}^1(\tau; j|i)$ that the system changes, for the first time, to the detected state j before time τ given that the preceding detected state was $i \neq j$ (cf. (13)). Here, we remember that $P(\mathcal{H}_n, \tau)$ is the probability that the path $\mathcal{H}_n$ is traversed before time τ, but not extended by additional states (cf. (49)). Since setting $w(\mathbf{j}) = 0$ guarantees that the final
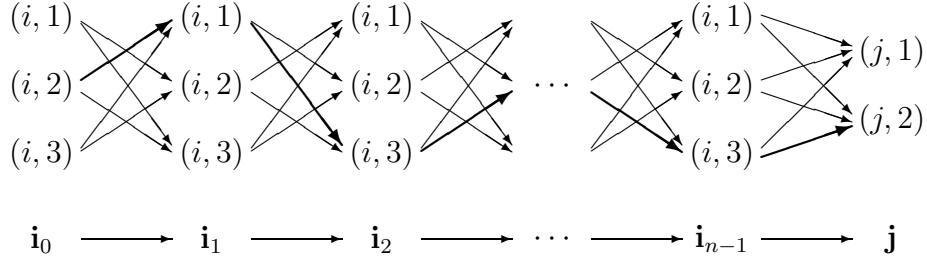
Figure 5: General hidden state model including parallel *and* sequential transitions illustrated for the case $N_i = 3$ and $N_j = 2$. The possible hidden transitions leading from the detected state $i \in \{(i,1),(i,2),(i,3)\}$ to the detected state $j \in \{(j,1),(j,2)\}$ are represented by arrows. The thick arrows indicate one of the numerous paths $\mathcal{H}_n$ of the form (82) which the system may take.

state $\mathbf{j}$ of $\mathcal{H}_n$ is not left any more, $P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0)$ is the probability that the path $\mathcal{H}_n$ is traversed before time τ. For this reason, $F_{\text{eff}}^1(\tau; j|i)$ is given by the sum over the probabilities $P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0)$ of all paths $\mathcal{H}_n$ which have the form (82):

$$F_{\text{eff}}^1(\tau; j|i) = \frac{1}{P(i)} \sum_{\mathcal{H}} P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0) = \frac{1}{P(i)} \sum_{n=0}^{\infty} \sum_{h_n=1}^{N_j} \sum_{\mathcal{H}_n} P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0) \,. \quad (84)$$

The factor $P(i)$ takes into account that $F_{\text{eff}}^1(\tau; j|i)$ is a conditional probability. It is determined by the probability

$$P(i) := \sum_{\substack{j=1 \\ (j \neq i)}}^{N} \sum_{n=0}^{\infty} \sum_{h_n=1}^{N_j} \sum_{\mathcal{H}_n} \lim_{\tau \to \infty} P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0) \quad (85)$$

that the system starts with the detected state i and changes to another state $j \neq i$ *at all*. It is obvious that the evaluation of formula (84) requires a computer program like EPIS.

From relation (84) we can obtain the *effective cumulative life-time distribution*

$$F_{\text{eff}}^1(\tau|i) := \sum_{\substack{j=1 \\ (j \neq i)}}^{N} F_{\text{eff}}^1(\tau; j|i) \quad (86)$$

analogous to (14) and confirm the desired *normalization condition*

$$\lim_{\tau \to \infty} F_{\text{eff}}^1(\tau|i) = 1 \,. \quad (87)$$

Moreover, we can derive the *effective survivor function*

$$S_{\text{eff}}^1(\tau|i) := 1 - F_{\text{eff}}^1(\tau|i) \,, \quad (88)$$

the *effective probability density function*

$$f_{\text{eff}}^1(\tau; j|i) := \frac{d}{d\tau} F_{\text{eff}}^1(\tau; j|i) \,, \quad (89)$$

and the *effective transition rates*

$$w_{\text{eff}}(\tau; j|i) := \frac{f_{\text{eff}}^1(\tau; j|i)}{S_{\text{eff}}^1(\tau|i)} \tag{90}$$

in accordance with (16) and (17). For the *effective overall departure rates*

$$w_{\text{eff}}(\tau|i) := \sum_{\substack{j=1 \\ (j \neq i)}}^{N} w_{\text{eff}}(\tau; j|i) \tag{91}$$

we again find the simple relation (67).

In their general form, the above formulas cannot be further evaluated. For illustrative reasons we will calculate the effective overall departure rates for the special case that there is exactly one hidden path of the form (82) which corresponds to the detected transition $j \longleftarrow i$. This implies that the length n of the path as well as the states $\mathbf{i}_k$ and $\mathbf{j}$ are uniquely determined by i and j, i.e. $n = n(i,j)$ and $h_k = h_k(i,j)$ ($k \in \{1,2,\ldots,n\}$). Therefore, the hidden states $\mathbf{i}_1, \ldots, \mathbf{i}_{n-1}$ could be interpreted as *transient states*. If all overall departure rates $w_k = w(i_k)$ are pairwise different from each other, we find

$$P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0)$$

$$= \left(\sum_{k=0}^{n-1} \frac{S^1(\tau|\mathbf{i}_k)}{[0 - w(\mathbf{i}_k)] \prod_{\substack{l=0 \\ (l \neq k)}}^{n-1} [w(\mathbf{i}_l) - w(\mathbf{i}_k)]} + \frac{1}{\prod_{\substack{l=0 \\ (l \neq n)}}^{n} [w(\mathbf{i}_l) - 0]} \right) w(\mathcal{H}_n) P(\mathbf{i}_0, \tau_0) \tag{92}$$

and, due to $w(\mathbf{i}_k) > 0$,

$$\lim_{\tau \to \infty} P(\mathcal{H}_n, \tau | w(\mathbf{j}) = 0) = \frac{1}{\prod_{l=0}^{n-1} w(\mathbf{i}_l)} w(\mathcal{H}_n) P(\mathbf{i}_0, \tau_0). \tag{93}$$

This finally yields

$$F_{\text{eff}}^1(\tau|i) = \frac{\left(\sum_{k=0}^{n-1} \dfrac{S^1(\tau|\mathbf{i}_k)}{-w(\mathbf{i}_k) \prod_{\substack{l=0 \\ (l \neq k)}}^{n-1} [w(\mathbf{i}_l) - w(\mathbf{i}_k)]} + \dfrac{1}{\prod_{l=0}^{n-1} w(\mathbf{i}_l)} \right) \sum_{\substack{j=1 \\ (j \neq i)}}^{N} w(\mathcal{H}_n) P(\mathbf{i}_0, \tau_0)}{\dfrac{1}{\prod_{l=0}^{n-1} w(\mathbf{i}_l)} \sum_{\substack{j=1 \\ (j \neq i)}}^{N} w(\mathcal{H}_n) P(\mathbf{i}_0, \tau_0)} \tag{94}$$

(which can be further simplified) and

$$S_{\text{eff}}^1(\tau|i) = \left(\prod_{l=0}^{n-1} w(\mathbf{i}_l) \right) \sum_{k=0}^{n-1} \frac{S^1(\tau|\mathbf{i}_k)}{w(\mathbf{i}_k) \prod_{\substack{l=0 \\ (l \neq k)}}^{n-1} [w(\mathbf{i}_l) - w(\mathbf{i}_k)]}. \tag{95}$$

$S^1_{\text{eff}}(\tau|i)$ is obviously a direct generalization of $S^1(\tau|i)$ since, for $n = 1$, we find

$$S^1_{\text{eff}}(\tau|i) = w(\mathbf{i}_0)\frac{S^1(\tau|\mathbf{i}_0)}{w(\mathbf{i}_0)} = S^1(\tau|\mathbf{i}_0) = S^1(\tau|i)\,. \tag{96}$$

Finally, we arrive at the desired result

$$w_{\text{eff}}(\tau|i) = \frac{\displaystyle\sum_{k=0}^{n-1} \frac{S^1(\tau|\mathbf{i}_k)}{\displaystyle\prod_{\substack{l=0 \\ (l\neq k)}}^{n-1} [w(\mathbf{i}_l) - w(\mathbf{i}_k)]}}{\displaystyle\sum_{k=0}^{n-1} \frac{S^1(\tau|\mathbf{i}_k)}{w(\mathbf{i}_k)\displaystyle\prod_{\substack{l=0 \\ (l\neq k)}}^{n-1} [w(\mathbf{i}_l) - w(\mathbf{i}_k)]}} \tag{97}$$

due to

$$\frac{d}{d\tau}S^1(\tau|\mathbf{i}_k) = -w(\mathbf{i}_k)S^1(\tau|\mathbf{i}_k)\,. \tag{98}$$

Although the special formula (97) is restricted to hidden state models consisting of sequences of n birth processes, it is already much more complicated than (68).

6 Summary and Outlook

In this paper we showed that, in the Markov case, survival analysis is related with the master equation which describes the temporal evolution of the distribution of states. The simulation of the master equation is a powerful technique for the investigation of various stochastically behaving systems in physics, chemistry, biology, and the social sciences. It is particularly suited for scenario techniques, since the numerical integration of the master equation is normally quite simple and fast.

However, some interesting questions related with the longitudinal time series of stochastic processes cannot be answered by means of the master equation. Whereas the microsimulation of sample time series can be done with the Monte-Carlo technique, it is rather inefficient for the numerical determination of quantities related with the sequencing of time series. Therefore, the simulation tool EPIS has recently been developed at the University of Stuttgart. It facilitates the generation of relevant paths and the evaluation of formulas which we were able to derive for path-related quantities. This includes the occurence probabilities and occurence time distributions of paths, or the expected escape time from undesired states.

Finally, the formula for the occurence probabilities of paths allowed to develop a hidden state concept of behavioral changes which can serve as a means for interpreting the respective time-dependence of hazard rates. Starting from a certain hidden state model, it is possible to derive the corresponding effective transition rates which can be compared with the time-dependence of the empirically obtained hazard rates. The different steps which are necessary for determining a suitable hidden state model and the corresponding parameter values (including the hazard rates) were illustrated by a concrete example.

Present research focuses on the investigation of the following questions:

1. Which kinds of time-dependences can be interpreted or approximated in terms of a hidden state model?

2. Does the time-dependence of hazard rates determine the corresponding hidden state model in a unique way?

3. If not, which are the transition schemes of the alternative hidden state models and which of them is the simplest or most plausible one in terms of a sociological or psychological interpretation?

Acknowledgements

References

Binder, K. (1979) *Monte Carlo Methods in Statistical Physics.* Berlin: Springer.

Blossfeld, H.-P., A. Hamerle, and K. U. Mayer (1986) *Ereignisanalyse.* Frankfurt: Campus.

Chiang, C. L. (1968) *Introduction to Stochastic Processes in Biostatistics.* New York: Wiley.

Courgeau, D. and É. Lelièvre (1992) *Event History Analysis in Demography.* Oxford: Clarendon Press.

Cox, D. R. (1975) "Partial likelihood", *Biometrica* **62**, 269–276.

Cox, D. R. and D. Oakes (1984) *Analysis of Survival Data.* London: Chapman and Hall.

Diekmann, A. and P. Mitter, eds. (1984) *Stochastic Modelling of Social Processes*, especially the contribution by A. Diekmann and P. Mitter: "A comparison of the 'sickle function' with alternative stochastic models of divorce rates", pp. 123–153. Orlando: Academic Press.

Elandt-Johnson, R. C. and N. L. Johnson (1980) *Survival Models and Data Analysis.* New York: Wiley.

Empacher, N. (1992) *Die Wegintegrallösung der Mastergleichung.* University of Stuttgart: PhD thesis.

Haken, H. (1983) *Synergetics,* 3rd edition. Berlin: Springer.

Helbing, D. (1994) "A contracted path integral solution of the discrete master equation". *Physics Letters A* **195**, 128–134.

Helbing, D. (1995) *Quantitative Sociodynamics. Stochastic Methods and Models of Social Interaction Processes.* Dordrecht: Kluwer Academic.

Helbing, D. and Molini, R. (1995) "Occurence probabilities of stochastic paths", *Physics Letters A* (submitted).

Kalbfleisch, J. D. and R. L. Prentice (1980) *The Statistical Analysis of Failure Time Data.* New York: Wiley.

Kinsey, A. C., W. B. Pomeroy, and C. E. Martin (1948) *Sexual Behavior in the Human Male.* Philadelphia: Saunders.

Lancaster, T. (1990) *The Econometric Analysis of Transition Data.* Cambridge: Cambridge University Press.

Molini, R. (1995) *Algorithmen und Anwendungen zur Pfadsummenlösungsmethode der Masterequation.* University of Stuttgart: Master's thesis.

Pearl, R. (1924) *Studies in Human Biology.* Baltimore: Williams and Wilkins.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992) *Numerical Recipes in C. The Art of Scientific Computing,* 2nd edition. Cambridge: Cambridge Univerisity Press.

Schildt, H. (1990) *C: The Complete Reference.* Berkeley: McGraw Hill.

Stewart, I. and D. Tall (1983) *Complex Analysis.* Cambridge: Cambridge University Press.

Troitsch, K. G. (1990) *Modellbildung und Simulation in den Sozialwissenschaften.* Opladen: Westdeutscher Verlag.

Tuma, N. B. and M. T. Hannan (1984) *Social Dynamics. Models and Methods.* Orlando: Academic Press.

Verhulst, P.-F. (1845) "Recherches mathématiques sur la loi d'accroissement de la population". *Nuoveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles* **18**, 1–41.

Weidlich, W. (1991) "Physics and social science—The approach of synergetics". *Physics Reports* **204**, 1–163.

Weidlich, W. and G. Haag (1983) *Concepts and Models of a Quantitative Sociology. The Dynamics of Interacting Populations.* Berlin: Springer.