ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

MARKOV STOCHASTIC PROCESSES IN BIOLOGY AND MATHEMATICS – THE SAME, AND YET DIFFERENT

MIŁOSŁAWA SOKÓŁ*

Abstract. Virtually every biological model utilising a random number generator is a Markov stochastic process. Numerical simulations of such processes are performed using stochastic or intensity matrices or kernels. Biologists, however, define stochastic processes in a slightly different way to how mathematicians typically do. A discrete-time discrete-value stochastic process may be defined by a function $p : X_0 \times X \to \{f : \Upsilon \to [0,1]\}$, where X is a set of states, X_0 is a bounded subset of X, Υ is a subset of integers (here associated with discrete time), where the function p satisfies $0 < p(x,y)(t) < 1$ and $\sum_y p(x,y)(t) = 1$. This definition generalizes a stochastic matrix. Although X_0 is bounded, X may include every possible state and is often infinite. By interrupting the process whenever the state transitions into the $X - X_0$ set, Markov stochastic processes defined this way may have non-quadratic stochastic matrices. Similar principle applies to intensity matrices, stochastic and intensity kernels resulting from considering many biological models as Markov stochastic processes. Class of such processes has important properties when considered from a point of view of theoretical mathematics. In particular, every process from this class may be simulated (hence they all exist in a physical sense) and has a well-defined probabilistic space associated with it.

Key words. stochastic matrix, stochastic kernel, intensity matrix, intensity kernel, simulation of stochastic process, probability space for Markow stochastic process

AMS subject classifications. 60J05, 60J10, 60J22, 60J25, 60J27, 60J45

1. Introduction.

Many branches of biology have recently realised the educational value of using modeling to represent the objects of their studies. Biological models include dynamical models showing a change of some biological quantity in time or the so-called models of artificial life [9],[11],[12]. Such models are physical in nature, i.e. they are experiments performed *in silico* using virtual objects to represent physical objects (at a scale ranging from genes, cells and organs to individuals, populations and whole ecosystems) in a programmed environment. As a result, they allow us to observe the dynamics of various biological phenomena. Despite their complexity, such models are only simplifications of real biological processes which utilise random number generators to simulate possible events in the system. A probability of an occurrence of a given event at given point in time depends on the features of the object and state of the environment at that point in time. Using mathematical nomenclature, they are stochastic processes with Markov property. In this study, mathematical properties of these models are discussed.

Numerical simulations of such models may be performed to obtain multiple realizations of the stochastic process. Using data gathered this way, it is possible to draw conclusions from the model with an aid of statistical methods. The disadvantage of this approach is that each simulation requires a substitution of numerical values for model parameters. If the parameters cannot be estimated in a reliable way, the above approach may not be entirely feasible. There therefore arises a necessity for mathematical analysis of the properties of this type of models.

Researchers using such models tend not to use mathematical terminology, perhaps due to the differences (a few minor and one major) in how biologists and mathematicians define stochastic processes. For a mathematician, a stochastic process is a collection of random variables, while for a biologist, it is a probability space with

*Faculty of Biology, Biological and Chemical Research Centre, University of Warsaw (milka@biol.uw.edu.pl).

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

functions as events, as summarised in Table 1 below.

TABLE 1
Mathematical and Biological definitions of stochastic processes

mathematical	Stochastic process is a collection of random variables defined on a common probability space $(\Omega, \sigma(\Omega), P)$, indexed by Υ which all take values in the same mathematical measureable space (X, Σ).
biological	Stochastic process as a probability space $(\Omega, \sigma(\Omega), P)$ in which the events are functions $\varphi : \Upsilon \to X$ taking values in a measureable space (X, Σ) where Σ is a σ-algebra generated by sets $\{\varphi(t); \varphi \in A \in \sigma(\Omega)\}$.

There exists a simple relationship between the mathematical nad biological definitions. For each probability space $(\Omega, \sigma(\Omega), P)$ whose events are functions $\varphi : \Upsilon \to X$, one may define a family of functions $(\xi_t)_{t\in\Upsilon}$ such that $\xi_t(\varphi) = \varphi(t)$, so that $(\xi_t)_{t\in\Upsilon}$ form a mathematical stochastic process. In this sense, the mathematical definition of a stochastic process is perhaps more general.

Biologists are in general aware of the Markov property and often treat it as an implicit feature of biological models. In such models, the probabilities of events forming the next state are calculated basing exclusively on the current state of the system. If we think of an indexing set as time time, then the mathematical and biological definitions of the Markov property are equivalent. But biologists do not generalize this property to filtrations, as they consider only very specific indexing sets: $\Upsilon = \{0, 1, \ldots, T\}$, $\Upsilon = \mathbb{N}_0$, $\Upsilon = [0, T]$ or $\Upsilon = [0, \infty)$.

To correctly analyse a biological model, it is crucial to identify those variables which allow creating a multidimensional equivalent of the stochastic matrix (for time-discrete models) or intensity matrix (for time-continuous models). For complicated models this may prove to be the most difficult part of mathematical analysis (this problem is not addressed in this article). Each artificial life model, however, has a collection of variables which form an associated stochastic process with Markov property, so that the conditional probabilities (probability rates) in stochastic (intensity) or matrix, may be expressed using functions with parameters. This is the first step in translating the biological model to the language of mathematics. For any well-defined stochastic or intensity matrix, there exists a Markov stochastic process whose realizations correspond to the model output. Analysis of such a Markov model is simpler than the analysis of the initial biological model.

2. Stochastic and intensity matrices and kernels. Due to diurnal or annual fluctuations of probabilities in many biological phenomena, a lot of biological models are non-homogeneous. But the true reason of starting the theory of Markov processes from non-homogeneous ones is the fact, that it is easier. Integral equations used to initialise simulation programs are simpler for a more general theory than its particular case. So, the items in stochastic or intensity matrix are time-dependent functions whose values are probabilities or probability rates.

Biologists usually consider states as more important than probabilities or probability rates. Hence they tend to display two-dimensional stochastic matrices transposed relative to how mathematicians would display them (Fig. 1) [14], [17]. Very often such matrices include also conditional expected value of the next state which allows for prediction of the features of a simulation.

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

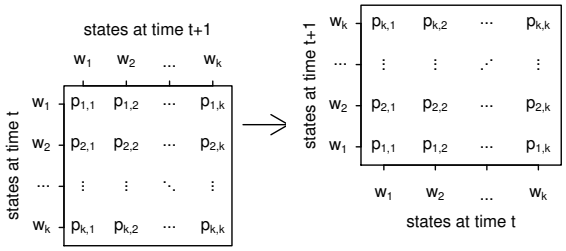

FIG. 1. *Stochastic matrix in traditional orientation (left) and functional orientation (right). The way the stochastic matrix is shown is determined by the nature of the states of the stochastic process (independent of the horizontal axis, but dependent on the vertical axis).*

Because the probabilities are calculated according to $p : X \times X \to \{f : \Upsilon \to [0,1]\}$, where X is a space of states, $\Upsilon \subset \mathbb{N}$, the stochastic matrix is very often infinite. No modeller, however, uses such matrices in practice. All simulators of biological phenomena impose limits on the number of initial states. Without this limitation, computer programs would be unable to handle the simulations. Simulation of such processes are nevertheless possible because whenever the subsequent state would leave the acceptable bounded set of states, the simulation is interrupted (even if the terminal time has not been reached). The set of target states does not have to be bounded . Hence modellers use non-square stochastic matrices. No biologist explicitly admits to using non-square matrices, but many of them do so.

The explicit consideration of non-square stochastic matrices allows for linear regression to be used for experiments where the conditional probabilities of transitions between the states are estimated using some independent variables (Fig. 2).

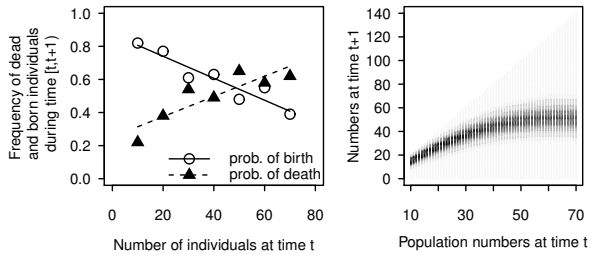

FIG. 2. *The regressions between number of individuals k at time t and the frequency births and deaths during time $[t, t+1)$ for $t \in [10, 70]$ (left). For $k < 10$ and $k > 70$ the probabilities of birth and death are impossible to foresee. The formulas $p_r(k)$ and $p_s(k)$ allow for calculating the probability of change of the population size from k to n during time $[t, t+1)$ using a formula:* $p(k,n)(t) = \sum_r \binom{k}{r}\binom{k}{k-n+r}(p_r(k))^r(1-p_r(k))^{k-r}(p_s(k))^{k-n+r}(1-p_s(k))^{n-r}$. *These probabilities are then inserted into a stochastic matrix (right). The darker the point, the higher the probability. The probabilities are positive for $k \in [10, 70]$ and $n \in [0, 2k]$. Minimal matrix which includes all positive probabilities has size $[10, 70] \times [0, 140]$.*

All finite square and not-square stochastic matrices $[x_{min}, x_{max}] \times [y_{min}, y_{max}]$ may be extended to half-infinite matrices by putting $p(x, y) = 0$ if $y \notin [y_{min}, y_{max}]$.

Let X be a subset of states, X_0 be a bounded subset of X, Σ be a σ-algebra defined on X, and $(\xi_t)_{t\in\Upsilon}$ be a stochastic process with probability space $(\Omega, \sigma(\Omega), P)$. Because the complicated models have a few variables forming stochastic process with

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

Markov property, it is better to accept a more functional definition of the stochastic matrix (Tab. 2).

TABLE 2
Mathematical and biological definitions of stochastic matrix for discrete-time discrete-value non-homogeneous Markov stochastic process $(\xi_t)_{t\in\Upsilon}$.

mathematical [13], [1], [3]	A stochastic (probability, transition, substitution) matrix is a collection of conditional probabilities: $p_{x,y}(t) = P\{\xi_{t+1} = y \mid \xi_t = x\}$ arranged in a matrix.
biological	A stochastic "matrix" is a function $p : X_0 \times X \to \{f : \Upsilon \to [0,1]\}$ and $p(x,y)(t) = P\{\xi_{t+1} = y \mid \xi_t = x\}$

The basic properties of this matrix are:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ 0 \le p_{x,y}(t) = p(x,y)(t) \le 1 \tag{2.1}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \sum_{y\in X} p_{x,y}(t) = \sum_{y\in X} p(x,y)(t) = 1 \tag{2.2}$$

The same applies to intensity matrices (Tab. 3). The basic properties of this

TABLE 3
Mathematical and biological definitions of intensity matrix for continuous-time discrete-value non-homogeneous Markov stochastic process $(\xi_t)_{t\in\Upsilon}$.

mathematical [6], [5]	An intensity (probability rate, transition rate) matrix is a collection of conditional probability rates: $q_{x,y}(t) = \lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta}=y \mid \xi_t=x\} - P\{\xi_t=y \mid \xi_t=x\}}{\Delta}$ arranged in a matrix.
biological	A intensity "matrix" is a function $q : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ and $q(x,y)(t) = \lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta}=y \mid \xi_t=x\} - P\{\xi_t=y \mid \xi_t=x\}}{\Delta}$

matrix are:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ q_{x,y}(t) = q(x,y)(t) \ge 0 \text{ if } x \ne y \tag{2.3}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ q_{x,x}(t) = q(x,x)(t) \le 0 \tag{2.4}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \sum_{y\in X} q_{x,y}(t) = \sum_{y\in X} q(x,y)(t) = 0 \tag{2.5}$$

In biological models, continuous variables appear just as frequently as discrete ones. This only makes sense if there exist a measure $\mu : \Sigma \to [0,\infty)$ on the subsets of X. For these variables slightly different definitions of (equivalents of) stochastic

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

TABLE 4
Mathematical and biological definitions of a stochastic kernel for discrete-time continuous-value non-homogeneous Markov stochastic process $(\xi_t)_{t\in\Upsilon}$.

mathematical	A stochastic kernel of a stochastic processes is a collection of densities of probability distributions p_t^x designated for all x and time t such that: $p_t^x : X \to [0,\infty)$ and $\int_U p_t^x(y)dy = P\{\xi_{t+1} \in U | \xi_t = x\}$ where U is a measurable subset of X.
biological	A stochastic kernel is a function $p : X_0 \times X \to \{f : \Upsilon \to [0,\infty)\}$ such that $\int_U p(x,y)(t)dy = P\{\xi_{t+1} \in U - \{x\} | \xi_t = x\}$ where U is a measurable subset of X.

or intensity matrix are needed. They are called stochastic kernels [10], [15] and for discrete time models the relevant definitions are shown in Tab. 4.

Note that mathematical and biological definitions are significantly different (not only due to different domains of the functions p). The basic properties of the mathematical stochastic kernels are:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ p_x(y)(t) \geq 0$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \int_X p_x(y)(t)dy = 1$$

The basic properties of the biological stochastic kernels are:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ p_{x,y}(t) \geq 0 \tag{2.6}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \int_X p(x,y)(t)dy \leq 1 \tag{2.7}$$

In many biological models probability of changing from state x to any state y ($y \neq x$) is very small. Many time-steps result in no change of the state x, so modelling it using a mathematical stochastic kernel may not be necessary. Using biological definition, a probability of no change of state x is equal to $1 - \int_y p(x,y)(t)dy$. If the change does not happen, then the next state may be determined from the distribution with density

$$p_x(t) : X \ni y \to \frac{p(x,y)(t)}{\int_X p(x,y)(t)dy} \in [0,\infty)$$

It is possible to formulate a definition of an intensity kernel (Tab.5). It is necessary for a Markov continuous-time continuous-value stochastic process. These processes appear in some biological models, but I was unable to find a mathematical article with definition of an intensity kernel although such Markov processes were discussed in mathematics [4], [2].

The basic properties of such intensity kernels are:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ q_{x,y}(t) \geq 0 \tag{2.8}$$

$$\exists_{M>0}\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \int_X q(x,y)(t)dy \leq M \tag{2.9}$$

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

TABLE 5
Mathematical and biological definitions of an intensity kernel for continuous-time continuous-value non-homogeneous Markov stochastic process $(\xi_t)_{t\in\Upsilon}$.

mathematical	???
biological	An intensity kernel is a function $q : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ such that $\int_U q(x,y)(t)dy = \lim_{\Delta\to 0} \frac{P\{\xi_{t+1}\in U-\{x\}|\xi_t=x\}}{\Delta}$, where U is a measurable subset of X.

Very often in biological models X is a subset of $\mathbb{F}^k \times \mathbb{N}^n \times \mathbb{R}^m$ (where $\mathbb{F}$ is some finite set). In such a case, in the second property above (and in the corresponding one for an intensity matrix) the integral would be replaced as a sum of integrals. As in statistics, when discussing basic theorems it is more convenient to consider discrete and continuous variables separately.

I will not prove that the biological definitions resulting from the analysis of biological models are correct, but I am will prove that for any function $p : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ with properties (1), (2) or (6), (7), and a function $q : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ with properties (3), (4), (5) or (8), (9), there exists a Markov stochastic process such that its stochastic or intensity matrix or kernel are equivalent to one of these functions. Such existence will be showing that it is possible to simulate such process. At the end, a construction of the probability space for such process will be presented.

3. Definitions. Let X be a set of states. Let $X_0 \subset X$ be bounded. Let $\Upsilon = \{0, 1, \ldots, T\}$ or $\Upsilon = \mathbb{N}_0$. For such sets there exist many functions satisfying conditions (1)-(9) for stochastic or intensity matrix or kernel. We will refer to such function as a stochastic function or an intensity function.

Definition 1. Let X be finite or countable. A stochastic discrete function is a function $p : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ satisfying the following conditions:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ 0 \le p(x,y)(t) \le 1 \tag{3.1}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \sum_{y\in X} p(x,y)(t) = 1 \tag{3.2}$$

Definition 2. Let X be finite or countable. An intensity discrete function is a function $q : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ satisfying the following conditions:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ q(x,y)(t) \ge 0 \text{ if } x \ne y \tag{3.3}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ q(x,x)(t) \le 0 \tag{3.4}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \sum_{y\in X} q(x,y)(t) = 0 \tag{3.5}$$

If X is uncountable, we will refer to the equivalents of stochastic or intensity kernels as continuous functions, although such functions need not to be continuous in the topological sense. Similarly as in statistics, a continuous variable need not to be continuous. However, I am not aware of any biological model with topologically discontinuous stochastic or intensity kernel.

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

Definition 3. Let X be uncountable. A stochastic continuous function is a function $p : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ satisfying the following conditions:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ p(x,y)(t) \geq 0 \tag{3.6}$$

$$\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \int_X p(x,y)(t)dy \leq 1 \tag{3.7}$$

Definition 4. Let X be uncountable. An intensity continuous function is a function $q : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ satisfying the following conditions:

$$\forall_{x\in X_0}\forall_{y\in X}\forall_{t\in\Upsilon}\ q(x,y)(t) \geq 0 \tag{3.8}$$

$$\exists_{M>0}\forall_{x\in X_0}\forall_{t\in\Upsilon}\ \int_X q(x,y)(t)dy \leq M \tag{3.9}$$

For all the functions defined above, physical existence of the corresponding stochastic Markov processes will be proved.

4. Simulation. By "physical existence of a Markov stochastic process" we understand that there exists an experiment which allows for simulation of a stochastic process with the Markov and other appropriate properties. Such an experiment may be programmed what allows for it multiply faster simulation than simulation of the original biological model.

There exist easy schemes for an experiment which simulates realizations of a stochastic process. Given an initial state x_0 at time $t = 0$ and a terminal time T, the next state at the next timestep is chosen according to a prescribed rule. The outcome of such an experiment may be thought of as a realization of a stochastic process. Examples of such schemes are shown in Fig. 3.

Description of the simulation of the process is reduced to an correct choice of the next transition time and the next state when the state at time t and the values of stochastic or intensity function are known. By "correct" we understand that the mean frequency of the choice of y as the state following state x must be equal to the formulas shown in biological definitions of stochastic and intensity matrices and kernels (i.e. Tables 2-4).

Theorem 1. Let $p : X_0 \times X \to \{f : \Upsilon \to [0,1]\}$ be a stochastic discrete function. For each state x at time t there exists a rule for choosing the next state y such that mean frequency of obtaining state y at time $t+1$ following state x at time t (probability P_t^x) is equal to $P_t^x(y) = p(x,y)(t)$.

Proof. An algorithm for simulating a discrete-value discrete-time stochastic process begins with ordering the states in space X in a sequence $(x_i)_{i=0,1,2,\ldots}$. Let x be a state at time t. A random number α is then drawn from the interval $[0,1]$. The next state y at time $t+1$ is equal to such x_i that:

$$\alpha \in \left[\sum_{j=0}^{i-1} p(x,x_j)(t), \sum_{j=0}^{i} p(x,x_j)(t)\right) \tag{4.1}$$

Such state always exists and the method for determining the state y induces a probabilistic space $(X, 2^X, P_t^x)$ such that:

$$P_t^x(x_i) = \sum_{j=0}^{i} p(x,x_j)(t) - \sum_{j=0}^{i-1} p(x,x_j)(t) = p(x,x_i)(t) \tag{4.2}$$

□

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

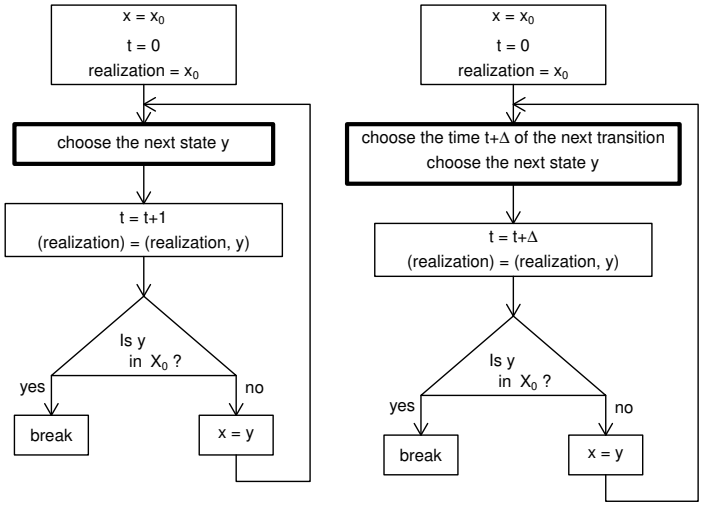

FIG. 3. *Algorithms for simulating discrete time (left) and continuous time (right) stochastic processes. By (realization)=(realization,y) we mean appending the next state to the list of previous states (in a discrete time case) or extending the realization to time t+Δ with state y (in a continuous time case).*

Theorem 2. Let $q : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}; f \text{ is integrable}\}$ be an intensity discrete function. For each state x at time t there exists a rule for choosing the next state y and the time $t + \Delta$ of transition from x to y such that the mean frequency of obtaining state y following a state x satisfies the following condition: $\lim_{\Delta\to 0} \frac{P_t^x(y)(t+\Delta)-P_t^x(y)(t)}{\Delta} = q(x, y)$, where $P_t^x(y)(t + \Delta)$ is the probability that the transition from x to y takes place at time $t + \Delta$.

Proof. An algorithm for simulating a continuous-time discrete-value stochastic process begins with ordering the states in space X in a sequence $(x_i)_{i=0,1,2,\ldots}$. Let $x = x_k$ be a state at time t. Because $q(x, x)(t) \leq 0$, the function

$$[0, T) \ni \Delta \to \exp\Big(\int_t^{t+\Delta} q(x, x)(s)ds\Big) \in \mathbb{R} \tag{4.3}$$

decreases from 1 to some $\theta \geq 0$. A random number α is then drawn from interval $[0, 1)$. If $\alpha \leq \theta$, then Δ becomes infinite and the system remains at state x until the terminal time is reached. Otherwise the time Δ is determined from the equation

$$\alpha = \exp\Big(\int_t^{t+\Delta} q(x, x)(s)ds\Big) \tag{4.4}$$

The time of the next transition is equal to $t + \Delta$. The next state is estimated by drawing a random number β from $[0, 1)$ and choosing such a state x_i that:

$$\beta \in \Big[\sum_{j=0, j\neq k}^{i-1} \frac{q(x, x_j)(t + \Delta)}{-q(x, x)(t + \Delta)}, \sum_{j=0, j\neq k}^{i} \frac{q(x, x_j)(t + \Delta)}{-q(x, x)(t + \Delta)}\Big) \tag{4.5}$$

The probability P_t^x determined using the above rules satisfies the condition:

$$\lim_{\Delta\to 0} \frac{P_t^x(x)(t + \Delta) - P_t^x(x)(t)}{\Delta} = \frac{P_t^x(x)(t + \Delta) - 1}{\Delta} = \tag{4.6}$$

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

$$= \lim_{\Delta\to 0} \frac{\exp\left(\int_t^{t+\Delta} q(x.x)(s)ds\right) - 1}{\Delta} = \frac{d}{d\Delta}\Big[\exp\Big(\int_t^{t+\Delta} q(x.x)(s)ds\Big)\Big] =$$

$$= \exp\Big(\int_t^t q(x.x)(s)ds\Big) \cdot \frac{d}{d\Delta}\int_t^{t+\Delta} q(x.x)(s)ds = q(x,x)(t)$$

For $y \neq x$ the $\frac{q(x,y)(t+\Delta)}{-q(x,x)(t+\Delta)}$ is a probability that the state x changes to state y under the condition that a significant change takes place. Hence:

$$\lim_{\Delta\to 0} \frac{P_t^x(y)(t+\Delta) - P_t^x(y)(t)}{\Delta} = \frac{P_t^x(y)(t+\Delta)}{\Delta} = \tag{4.7}$$

$$= \lim_{\Delta\to 0} \frac{1}{\Delta} \frac{q(x,y)(t+\Delta)}{-q(x,x)(t+\Delta)}\Big(1 - P_t^x(x)(t+\Delta)\Big) =$$

$$= \lim_{\Delta\to 0} \frac{q(x,y)(t+\Delta)}{-q(x,x)(t+\Delta)} \lim_{\Delta\to 0} \frac{1 - P_t^x(x)(t+\Delta)}{\Delta} =$$

$$= \frac{q(x,y)(t)}{-q(x,x)(t)}(-q(x,x)(t)) = q(x,y)$$

∎

A simulaton of discrete-time continuous-value stochastic process requires an algorithm to draw random numbers from a given distribution. A small, but effective and universal physical algorithm is:

1. draw a point from a rectangle containing a large enough part of the graph of the density of this distribution,
2. if the point lies above the graph of density, reject it and repeat step 1.,
3. read off the first coordinate of the first point which was not rejected in step 2.

Even though the above algorithm is effective, for many distributions there exist more robust algorithms to draw appropriate random numbers [7], [8].

Theorem 3. Let $p : X_0 \times X \to \{f : \Upsilon \to [0,\infty)\}$ be a stochastic continuous function. For each state x at time t there exists a rule of choosing the next state y such that the mean frequency of obtaining a state from $U \in \Sigma$ at time $t+1$ following state x at time t (probability P_t^x) is equal $P_t^x(U) = \int_U p(x,y)(t)dy$.

Proof. Let x be a state at time t. Then the integral $\int_X p(x,y)(t)dy$ needs to be calculated. The number α is drawn from the interval $[0,1)$. If $\alpha > \int_X p(x,y)(t)dy$, then the next state y is equal to x. If $\alpha \leq \int_X p(x,y)(t)dy$ then a state y is drawn from the distribution with density:

$$y \to \frac{p(x,y)(t)}{\int_X p(x,y)(t)dy} \tag{4.8}$$

The above operations induce a probability space (X, Σ, P_t^x) which may be broke down as follows: first it is determined whether the state x undergoes a change or not (i.e. Bernoulli distribution) and if it does, the distribution of the next state has density 4.8. In this probability space, for any $U \in \Sigma$, we have:

$$P_t^x(U) = P_t^x(U|x \text{ changes to } y \in U)P_t^x(x \text{ changes to } y \in U) = \tag{4.9}$$

$$= \frac{\int_U p(x,y)(t)dy}{\int_X p(x,y)(t)dy} \int_X p(x,y)(t)dy = \int_U p(x,y)(t)dy$$

∎

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

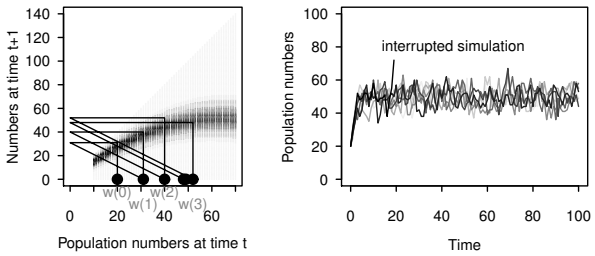

FIG. 4. *Left: stochastic matrix calculated in Fig. 2 and a sample realization* $(w(0), w(1), \dots)$ *of the process. Right: a graph with five sample realizations of the process over 100 timesteps.*

A simulation of a stochastic process which uses two-dimensional stochastic functions (stochastic matrix or kernel) has a clear graphical interpretation (Fig. 4).

Theorem 4. Let $q : X_0 \times X \to \{f : \Upsilon \to [0, \infty); f \text{ is integrable}\}$ be a continuous intensity function. For each state x at time t there exists a rule for choosing the next state y and time of transition $t+\Delta$ such that the mean frequency of obtaining a state from $U \in \Sigma$ following state x at time t satisfies the condition: $lim_{\Delta\to 0}\frac{P_t^x(U)(t+\Delta)}{\Delta} = \int_U q(x,y)(t)dy$, where $P_t^x(U)(t+\Delta)$ denotes the probability that the transition from x to any $y \in U$ takes place at time $t+\Delta$.

Proof. Let x be a state at time t. Let $q_{x,x} : [0,T) \ni s \to \int_X q(x,y)(s)dy$. The function $\Delta \to \exp\Big(-\int_t^{t+\Delta} q_{x,x}(s)ds\Big)$ decreases from 1 to some $\theta \geq 0$. If $\alpha \leq \theta$, then Δ becomes infinite and the system remains at state x until the terminal time is reached. Otherwise the time Δ is determined from the equation:

$$\alpha = \exp\Big(-\int_t^{t+\Delta} q_{x,x}(s)ds\Big) \tag{4.10}$$

The time of the next transition is equal to $t+\Delta$. The next state y is estimated by the drawing a random number from the distribution with density $y \to \frac{q(x,y)(t+\Delta)}{q_{x,x}(t+\Delta)}$.

A probability space (X, Σ, P_t^x) formed by the above rules satisfies the condition:

$$lim_{\Delta\to 0}\frac{P_t^x(U)(t+\Delta)}{\Delta} = \tag{4.11}$$

$$= lim_{\Delta\to 0}\frac{1}{\Delta}\frac{\int_U q(x,y)(t+\Delta)dy}{q_{x,x}(t+\Delta)}P_t^x\{x \text{ changes to any state at } t+\Delta\} =$$

$$= lim_{\Delta\to 0}\frac{1}{\Delta}\frac{\int_U q(x,y)(t+\Delta)dy}{q_{x,x}(t+\Delta)}\big(1 - P_t^x\{x \text{ does not change in } [t,t+\Delta)\}\big) =$$

$$= lim_{\Delta\to 0}\frac{1}{\Delta}\frac{\int_U q(x,y)(t+\Delta)dy}{q_{x,x}(t+\Delta)}\Big(1 - \exp\Big(-\int_t^{t+\Delta} q_{x,x}(s)ds\Big)\Big) =$$

$$= \frac{\int_U q(x,y)(t)dy}{q_{x,x}(t)} lim_{\Delta\to 0}\frac{\Big(1 - \exp\Big(-\int_t^{t+\Delta} q_{x,x}(s)ds\Big)\Big)}{\Delta} =$$

$$= -\frac{\int_U q(x,y)(t)dy}{q_{x,x}(t)}\frac{d}{d\Delta}\exp\Big(-\int_t^{t+\Delta} q_{x,x}(s)ds\Big) =$$

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

$$= -\frac{\int_U q(x,y)(t)dy}{q_{x,x}(t)} \exp\Big(-\int_t^t q_{x,x}(s)ds\Big)\frac{d}{d\Delta}\Big(-\int_t^{t+\Delta} q_{x,x}(s)ds\Big) =$$

$$= -\frac{\int_U q(x,y)(t)dy}{q_{x,x}(t)}(-q_{x,x}(t)) = \int_U q(x,y)(t)dy \qquad \square$$

If there exists a probability space $(\Omega, \sigma(\Omega), P)$ on the set of all realizations of the simulations, then for any state x and time t: $P\{\varphi \in \Omega; \varphi(t) = x\} = 0$ for continuous stochastic functions and continuous intensity functions. Therefore the probabilities P_t^x are defined inside a set of measure 0. It is sensible, similarly as it is sensible to define length or area for one- or two-dimensional geometric objects in a three-dimensional space.

It remains to be discussed whether the probabilistic space $(\Omega, \sigma(\Omega), P)$ exists.

5. Probability space for limited time. The classical theorem about the probability space for a stochastic process induced by finite square stochastic and intensity matrices goes back to Kolmogorov in as early as late 40s. But for some infinite matrices, the sum of probabilities over all trajectories is less than 1. But what may be said about half-infinite matrices?

In this section I will show that there exists a probability space $(\Omega, \sigma(\Omega), P)$ for any stochastic or intensity function defined in chapter 3 such that:

1. Ω is a set of all realizations of the simulations described in chapter 4,
2. $\sigma(\Omega)$ is a sigma-algebra on Ω,
3. P is a probability that a specified realization appears during simulation.

Therefore the stochastic processes will be considered more from a biologist's than a mathematician's perspective.

Theorem 5. Let X be a countable set of states. Let X_0 be finite. Let T be the (finite) terminal time. Let $x_0 \in X_0$. Let $p : X_0 \times X \to f : 0, 1, \ldots, T \in [0, 1]$ be a discrete stochastic function. Then a structure $(\Omega, \sigma(\Omega), P)$ such that:

1. $\Omega = \{(x_0, x_1, ..., x_{t_m}); x_0, x_1, ..., x_{t_m-1} \in X_0$ and if $t_m < T$ then $x_{t_m} \in X - X_0\}$,
2. $\sigma(\Omega) = 2^\Omega$,
3. if $a = (x_0, x_1, ..., x_{t_m})$ is a realization of a simulation, then $P(a) = \prod_{i=1}^{t_m} p(x_{i-1}, x_i)$,
4. if $A \subset \Omega$ then $P(A) = \sum_{a \in A} P(a)$,

is a well-defined probability space.

Proof. The set Ω is countable, so P is a measure. It is necessary to prove that $P(\Omega) = 1$. But for any T we have:

$$\sum_{t_m=1}^{T-1} \sum_{x_1 \in X_0} \cdots \sum_{x_{t_m-1} \in X_0} \sum_{x_{t_m} \in X - X_0} p(x_0, x_1)(0)...p(x_{t_m-1}, x_{t_m})(t_m - 1) +$$

$$+ \sum_{x_1 \in X_0} \cdots \sum_{x_{T-1} \in X_0} \sum_{x_T \in X} p(x_0, x_1)(0)...p(x_{T-1}, x_T)(T-1) = 1 \tag{5.1}$$

We will use mathematical induction over the terminal time T. If $T = 1$, then realizations are precisely the pairs (x_0, x), where x is any state from X. Because p is a stochastic function:

$$P(\Omega) = \sum_{x_1 \in X} p(x_0, x_1) = 1 \tag{5.2}$$

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

If the assumption is true up to time T, then for $T+1$:

$$P(\Omega) = \sum_{t_m=1}^{T} \sum_{x_1\in X_0} \cdots \sum_{x_{t_m-1}\in X_0} \sum_{x_{t_m}\in X-X_0} p(x_0,x_1)(0)...p(x_{t_m-1},x_{t_m})(t_m-1)+$$

$$+ \sum_{x_1\in X_0} \cdots \sum_{x_T\in X_0} \sum_{x_{T+1}\in X} p(x_0,x_1)(0)...p(x_T,x_{T+1})(T) = \tag{5.3}$$

$$= \sum_{t_m=1}^{T-1} \sum_{x_1\in X_0} \cdots \sum_{x_{t_m-1}\in X_0} \sum_{x_{t_m}\in X-X_0} p(x_0,x_1)(0)...p(x_{t_m-1},x_{t_m})(t_m-1)+$$

$$+ \sum_{x_1\in X_0} \cdots \sum_{x_{T-1}\in X_0} \sum_{x_T\in X-X_0} p(x_0,x_1)(0)...p(x_{T-1},x_T)(T-1)+$$

$$+ \sum_{x_1\in X_0} \cdots \sum_{x_T\in X_0} p(x_0,x_1)(0)...p(x_{T-1},x_T)(T-1) \sum_{x_{T+1}\in X} p(x_T,x_{T+1})(T) =$$

$$= \sum_{t_m=1}^{T-1} \sum_{x_1\in X_0} \cdots \sum_{x_{t_m-1}\in X_0} \sum_{x_{t_m}\in X-X_0} p(x_0,x_1)(0)...p(x_{t_m-1},x_{t_m})(t_m-1)+$$

$$+ \sum_{x_1\in X_0} \cdots \sum_{x_{T-1}\in X_0} \sum_{x_T\in X-X_0} p(x_0,x_1)(0)...p(x_{T-1},x_T)(T-1)+$$

$$+ \sum_{x_1\in X_0} \cdots \sum_{x_T\in X_0} p(x_0,x_1)(0)...p(x_{T-1},x_T)(T-1) =$$

$$= \sum_{t_m=1}^{T-1} \sum_{x_1\in X_0} \cdots \sum_{x_{t_m-1}\in X_0} \sum_{x_{t_m}\in X-X_0} p(x_0,x_1)(0)...p(x_{t_m-1},x_{t_m})(t_m-1)+$$

$$+ \sum_{x_1\in X_0} \cdots \sum_{x_{T-1}\in X_0} \sum_{x_T\in X} p(x_0,x_1)(0)...p(x_{T-1},x_T)(T-1) = 1$$

according to the inductive assumption. So $P(\Omega) = 1$ for any finite time T. □

This theorem cannot be extended to infinite time, which is unfortunate, as the infinite time case often appears in biology. The probability space for infinite time will be discussed in the next chapter.

For a discrete intensity function, a realization is a step-function and it is often denoted by $x_0 \overset{t_1}{\to} x_1 \overset{t_2}{\to} ... \overset{t_m}{\to} x_m$, where t_i is a time of transition from x_{i-1} to x_i. Note that $x_0, x_1, ..., x_{m-1} \in X_0$ always, and if t_m¡T, then $x_m \in X - X_0$. The set of realizations Ω consists of all such functions. Construction of the probability space is similar to a construction of the Lebesgue measure in $\mathbb{R}^n$. First, the base sets of realisations (equivalents of n-dimensional rectangles) will be defined.

Definition 5. A base set $B^{x_0,x_1,...,x_m}_{\Delta_1,...,\Delta_m}$ is a subset of Ω determined by a sequence of states $(x_0, x_1, , x_m)$ and sequence of half-open intervals $(\Delta_1, \Delta_2, , \Delta_m)$ included in [0,T] such that:

1. $x_0, x_1, ..., x_{m-1} \in X_0$ and $x_m \in X$,
2. $\forall_i \forall_{t\in\Delta_i} \forall_{s\in\Delta_{i+1}} t < s$,
3. $\bigcup_i \Delta_i = [0, t_m)$,
4. if $t_m < T$ then $x_m \in X - X_0$,

Then $B^{x_0,x_1,...,x_m}_{\Delta_1,...,\Delta_m} = \{x_0 \overset{t_1}{\to} x_1 \overset{t_2}{\to} ... \overset{t_m}{\to} x_m; \forall_{t_i} t_i \in \Delta_i\}$

Base sets are sets of realizations which differ in transition times between states (Fig.5).

A function $\wp$ (equivalent to n-dimensional volume of n-dimensional rectangles) is defined over the base sets.

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

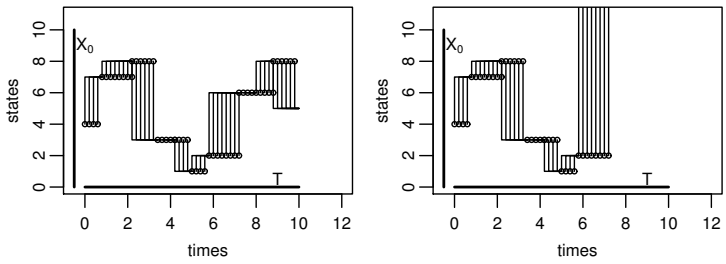

FIG. 5. *The idea of base sets for an intensity discrete functions.*

Definition 6. $\wp(B^{x_0,x_1,\ldots,x_m}_{\Delta_1,\ldots,\Delta_m}) = \prod_{i=1}^{m} \wp(B^{x_{i-1},x_i}_{\Delta_i})$ and when $\Delta_i = [t_{i-1}, t_i)$:

1. if $x_{i-1} = x_i = x$ then:

$$\wp\big(B^{x,x}_{[t_{i-1},t_i)}\big) = \exp\Big(\int_{t_{i-1}}^{t_i} q(x,x)(s)ds\Big) \tag{5.4}$$

2. if $x_{i-1} \neq x_i$ and $x_i \in X_0$ then:

$$\wp\big(B^{x_i,x_j}_{[t_{i-1},t_i)}\big) = \tag{5.5}$$

$$= \int_{t_{i-1}}^{t_i} q(x_{i-1},x_i)(t) \exp\Big(\int_{t_{i-1}}^{t} q(x_{i-1},x_{i-1})(s)ds + \int_{t}^{t_i} q(x_i,x_i)(s)ds\Big)dt$$

3. if $x_{i-1} \neq x_i$ and $x_i \in X - X_0$ then:

$$\wp\big(B^{x_i,x_j}_{[t_{i-1},t_i)}\big) = \int_{t_{i-1}}^{t_i} q(x_{i-1},x_i)(t) \exp\Big(\int_{t_{i-1}}^{t} q(x_{i-1},x_{i-1})(s)ds\Big)dt \tag{5.6}$$

These formulas do not come from nowhere. They are motivated by the approximation of the intensity function by the stochastic functions with decreasing timesteps and the necessity to calculate the probability of the occurrence of a given realization as a product of the suitable values of the stochastic functions. From these products, in a limiting case, we obtain precisely the formulas presented above.

It can be proved that such functions are well-defined and that they are countably additive. It can be proved that all assumptions of the Caratheodory theorem in the measure theory are satisfied. Hence an outer measure $\wp^*$ on 2^Ω such that for any $A \subseteq \Omega$:

$$\wp^*(A) = \inf_{\bigcup B_n \supseteq A} \sum \wp(B_n) \text{ where } B_n \text{ is base set}$$

may be defined. The outer measure limited to sets $A \subseteq \Omega$ satisfying Caratheodory condition:

$$\forall_{C \subseteq \Omega} \wp^*(A - C) + \wp^*(A \cap C) = \wp^*(A)$$

is a measure. But it may be proved that $\wp^*(\Omega) = 1$, so it is in fact a probability. Proofs of all theorems mentioned above were already published [16] and the details are omitted from this work for clarity. The constraints imposed on X_0 and time T are of significance when proving that $\wp^*(\Omega) = 1$.

Constructions of probability spaces for continuous stochastic and intensity functions are analogous to constructions of probability spaces for discrete and intensity

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

functions. But using two types of integral functions (on $\mathbb{R}$ and on X with different variables used as abscissa) results in proofs of poor readability. I believe a better strategy is to associate with the set of all possible realizations described in chapter 4 the sequence of ascending probabilistic spaces formed by discrete stochastic functions. The resulting corresponding theorems could read as follows:

Theorem 6. Let (X, Σ, μ) be a set of states with measure $\mu : \Sigma \to [0, \infty)$. Let X_0 be a bounded subset of X. Let $p : X_0 \times X \to \{f : \Upsilon \to [0, \infty); f \text{ is integrable }\}$ be a stochastic continuous function. For any n there exists a sequence $(A_1^n, A_2^n, ...)$ of measureable sets on X such that:

a1. $\forall_{i,j;i\neq j} A_i^n \cap A_j^n = \emptyset$,

a2. $\bigcup_{i=1}^{\infty} A_i^n = X$ and $\bigcup_{i=1}^{k_n} A_i^n = X_0$,

a3. $\forall_{A_i^{n+1}} \exists_{A_j^n} A_i^{n+1} \subset A_j^n$,

a4. $\forall_n \forall_i \mu(A_i^n) < \frac{1}{2^n}$.

If $X_n = \{A_1^n, A_2^n, ...\}$ and $X_{n,0} = \{A_1^n, A_2^n, ..., A_{k_n}^n\}$, and $p_n : X_{n,0} \times X_n \to \{f : \Upsilon \to [0, \infty)\}$, and $p_n(A_i^n, A_j^n)(t) = \frac{1}{\mu(A_i^n)} \int_{A_i} \int_{A_j} p(x, y)(t) dydx$ where $i \neq j$, and $p_n(A_i^n, A_i^n)(t) = \frac{1}{\mu(A_i^n)} \int_{A_i} \int_{A_i} p(x, y)(t) dydx + 1 - \frac{1}{\mu(A_i^n)} \int_{A_i} \int_X p(x, y)(t) dydx$, then p_n is a discrete stochastic function.

If $(\Omega_n, \sigma(\Omega_n), P_n)$ is a probability space for the stochastic discrete function p_n, then the structure $(\Omega, \sigma(\Omega), P)$ such that

t1. Ω is a set of all sequences $x_0 \to x_1 \to ... \to x_{t_m}$ such that $x_0, x_1, ...x_{t_m-1} \in X_0$ and if $t_m < T$ then $x_{t_m} \in X - X_0$,

t2. if $\gamma_n(A_0^n \to A_2^n \to ... \to A_{t_m}^n) = \{x_0 \to x_1 \to ... \to x_{t_m} \in \Omega; x_i \in A_i^n\}$ and $\forall_{U \subset \Omega} \gamma_n^{-1}(U) = \{A_0^n \to A_2^n \to ... \to A_{t_m}^n, \gamma_n(A_0^n \to A_2^n \to ... \to A_{t_m}^n) \in U\}$ and $\gamma_n^{-1}(\sigma(\Omega_n)) = \{\sigma_n^{-1}(U); U \in \sigma(\Omega_n)\}$ then $\sigma(\Omega) = \bigcup_n \gamma_n^{-1}(\sigma(\Omega_n))$,

t3. $\forall_{U \in \sigma(\Omega)} P(U) = \lim_{n\to\infty} P_n(\gamma_n^{-1}(U))$,

is a probabilistic space.

Theorem 7. Let (X, Σ, μ) be a set of states with measure $\mu : \Sigma \to [0, \infty)$. Let X_0 be a bounded subset of X. Let $p : X_0 \times X \to \{f : \Upsilon \to [0, \infty); f \text{ is integrable }\}$ be an intensity continuous function. For any n there exist a sequence $(A_1^n, A_2^n, ...)$ of measureable sets on X such that:

a1. $\forall_{i,j;i\neq j} A_i^n \cap A_j^n = \emptyset$,

a2. $\bigcup_{i=1}^{\infty} A_i^n = X$ and $\bigcup_{i=1}^{k_n} A_i^n = X_0$,

a3. $\forall_{A_i^{n+1}} \exists_{A_j^n} A_i^{n+1} \subset A_j^n$,

a4. $\forall_n \forall_i \mu(A_i^n) < \frac{1}{2^n}$.

If $X_n = \{A_1^n, A_2^n, ...\}$ and $X_{n,0} = \{A_1^n, A_2^n, ..., A_{k_n}^n\}$, and $p_n : X_{n,0} \times X_n \to \{f : \Upsilon \to [0, \infty)\}$, and $p_n(A_i^n, A_j^n)(t) = \frac{1}{\mu(A_i^n)} \int_{A_i} \int_{A_j} p(x, y)(t) dydx$ where $i \neq j$, and $p_n(A_i^n, A_i^n)(t) = \frac{1}{\mu(A_i^n)} \int_{A_i} \int_{A_i} p(x, y)(t) dydx - \frac{1}{\mu(A_i^n)} \int_{A_i} \int_X p(x, y)(t) dydx$, then p_n is an intensity discrete function.

If $(\Omega_n, \sigma(\Omega_n), P_n)$ is a probability space for the discrete intensity function p_n, then a structure $(\Omega, \sigma(\Omega), P)$ such that:

t1. Ω is a set of all step functions $x_0 \overset{t_1}{\to} x_1 \overset{t_2}{\to} ... \overset{t_m}{\to} x_m$ such that $x_0, x_1, ...x_m \in X_0$ and if $t_m < T$ then $x_m \in X - X_0$,

t2. if $\gamma_n(A_0^n \overset{t_1}{\to} A_1^n \overset{t_2}{\to} ... \overset{t_m}{\to} A_m^n) = \{x_0 \overset{t_1}{\to} x_1 \overset{t_2}{\to} ... \overset{t_m}{\to} x_m \in \Omega; x_i \in A_i^n\}$ and $\forall_{U \subset \Omega} \gamma_n^{-1}(U) = \{A_0^n \overset{t_1}{\to} A_1^n \overset{t_2}{\to} ... \overset{t_m}{\to} A_m^n, \gamma_n(A_0^n \overset{t_1}{\to} A_1^n \overset{t_2}{\to} ... \overset{t_m}{\to} A_m^n) \in U\}$ and $\gamma_n^{-1}(\sigma(\Omega_n)) = \{\sigma_n^{-1}(U); U \in \sigma(\Omega_n)\}$ then $\sigma(\Omega) = \bigcup_n \gamma_n^{-1}\sigma(\Omega_n)$,

t3. $\forall_{U \in \sigma(\Omega)} P(U) = \lim_{n\to\infty} P_n(\gamma_n^{-1}(U))$,

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

is a probabilistic space.

Both theorems are extensive because they assume and demonstrate a lot of simpler propositions. Both theorems are very similar, as they are based on the same intuition of transferring theorems about discrete distributions to continuous ones.

6. Probability space for infinity time. For stochastic and intensity functions, probabilistic spaces for $T = \infty$ must be different than those described in chapter 5. It is clear for stochastic discrete functions: suppose $p : X_0 \times X \to \{f :\to [0,1]\}$ is such that $p(x,y)(t) = 0$ if $y \in X - X_0$ and for all $x, y \in X_0$ it is always that $p(x,y)(t) < 1 - \epsilon$. Then $P(x_0 \to x_1 \to ...) = \prod_{i=1}^{\infty} p(x_{i-1}, x_i) = 0$ for all realizations in infinite time. Hence P defined as such a product is not a probability.

Let Ω_{t_m} be a set of all sequences or step functions such that the last state at time t_m is included $X - X_0$. Let Ω_∞ be a set of all sequences $\mathbb{N} \to X_0$ or all step functions $[0,\infty) \to X_0$. Let $\Omega = \Omega_{t_m} \cup \Omega_\infty$.

For any stochastic or intensity function let $(\Omega_T, \sigma(\Omega_T), P_T)$ be a probability space formed for a finite time T. Let $\gamma_T : \Omega \ni \varphi \to \varphi|_{[0,T]} \in \Omega_T$. Then a set of subsets of Ω given by $\{\gamma^{-1}(U); U \in \sigma(\Omega_T)\}$ is a σ-algebra on Ω, denoted by $\sigma_T(\Omega)$. It is generated by sets consisting of $x_0 \to x_1 \to ... \to x_{t_m}$ or $B^{x_0,x_1,...x_m}_{\Delta_1,\Delta_2,...,\Delta_m}$ and all possible sequences or step functions defined for times greater than t_m or $\sum \Delta_i$ (Fig.6).

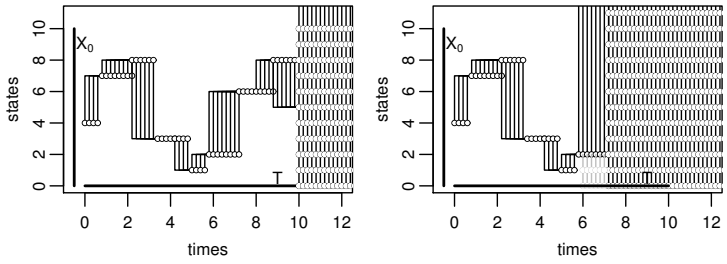

FIG. 6. *The idea of base sets in* $\sigma_T(\Omega)$.

Let $T_1 < T_2 < ...$ be a countable, unbounded sequence of times. Then $\sigma_{T_1}(\Omega) \subseteq \sigma_{T_2}(\Omega) \subseteq ....$ So a probability space $(\Omega, \sigma(\Omega), P)$ may be defined as follows:

1. $\sigma(\Omega) = \bigcup_n \sigma_{T_N}(\Omega)$,
2. $\forall_{A \in \sigma(\Omega)} \exists_n A \in \sigma_{T_N}(\Omega)$ and then $P(A) = P_{T_n}(A)$.

This probability space is well-defined.

In all previous considerations, the starting time for all simulations was $t = 0$ and initial state was x_0. But all the proofs were independent of the choices of initial time and state. In fact the initial time and state may be chosen to be any numbers from $\mathbb{R}$ and X_0 respectively. Probability space for initial time T_s, initial state x_0 and terminal time T_f will be denoted as $(\Omega^{x_0}_{[T_s,T_f]}, \sigma(\Omega^{x_0}_{[T_s,T_f]}), P^{x_0}_{[T_s,T_f]})$.

Let $x \in X_0$ and $t \in [T_s, T_f)$. The set $\{\varphi \in \sigma(\Omega^{x_0}_{[T_s,T_f]}); \varphi(t) = x\}$ will be denoted as A^x_t. Then $P_{[T_s,T_f)}(A^x_t) = P_{[T_s,t)}(A^x_t)$ and generally, if

$$A^{x_1,x_2,...,x_n}_{t_1,t_2,...,t_n} = \{\varphi \in \sigma(\Omega^{x_0}_{[T_s,T_f)}); \varphi(t_1) = x_1 \text{ and } ... \text{ and } \varphi(t_n) = x_n\} \tag{6.1}$$

then

$$P^{x_0}_{T_s,T_f}(A^{x_1,x_2,...,x_n}_{t_1,t_2,...,t_n}) = P^{x_0}_{T_s,t_1}(A^{x_1}_{t_1}) P^{x_1}_{t_1,t_2}(A^{x_2}_{t_2}) ... P^{x_{n-1}}_{t_{n-1},t_n}(A^{x_n}_{t_n}) \tag{6.2}$$

This is a result of defining a probability of any realization as a product of probabilities given by stochastic functions from successive time intervals.

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

7. Probability space for unbounded X_0. The ideas applied in the past chapter may be used to construct a probability space for a stochastic function $p : X \times X \to \{f : \Upsilon \to \mathbb{R}\}$ and intensity function $q : X \times X \to \{f : \Upsilon \to \mathbb{R}\}$.

Let Ω be the set of all sequences or all step functions defined on $\mathbb{N} \times X$ or $[T_s, \infty) \times X$. It consists of all infinite sequences or step functions, as well as sequences converging to infinity and all step functions converging to infinity in finite time or step functions wchich any subset of values converging to infinity in finite time (and, more general, all step function which subset of value converging to infinity or minus infinity in finite time).

Let $X_0 \subseteq X_1 \subseteq X_2 ...$ be bounded subsets of X such that $\bigcup_n X_n = X$. For all there X_n exists a probability space $(\Omega_n, \sigma(\Omega_n), P_n)$ defined by stochastic or intensity function whose domain is $X_n \times X$. Let $m_n(\varphi) = \min\{t \in \Upsilon; \varphi(t) \in X - X_n\}$ for $\varphi \in \Omega_n$. Let $\gamma_T : \Omega \ni \varphi \to \varphi|_{[0,m_n(\varphi)]} \in \Omega_n$. Then we may define a σ-algebra on Ω as $\sigma_n(\Omega) = \{\gamma_n^{-1}(A); A \in \sigma(\Omega_n)\}$.

These σ-algebras satisfy conditions: $\sigma_1(\Omega) \subseteq \sigma_2(\omega) \subseteq ....$ Furthermore their union $\sigma(\Omega) = \bigcup \sigma_n(\Omega)$ is a σ-algebra on Ω. A function $P : \sigma(\Omega) \to [0, 1]$ such that: if $A \in \sigma_n(\Omega)$ then $P(A) = P_n(A)$, is a probability. Thus $(\Omega, \sigma(\Omega), P)$ is a probability space for stochastic or intensity function $X \times X \to f : \Upsilon \to \mathbb{R}$. It differs from the mathematical probability space $(\Omega', \sigma(\Omega'), P')$ of a Markov stochastic process (if it exists) only in the set of events. The set Ω is larger than Ω'.

8. Stochastic process defined by a stochastic and an intensity functions. If we set $\xi_t : \Omega \ni \varphi \to \varphi(t) \in X$, then ξ_t is a random variable and $(\xi_t)_{t\in\Upsilon}$ is a stochastic process with probability space $(\Omega, \sigma(\Omega), P)$. This is a stochastic process defined by stochastic or intensity functions $X_0 \times X \to \{f : \Upsilon \to \mathbb{R}\}$. This stochastic process has the Markov property. This is a simple conclusion of equation (6.2).

To finish the proof of the general theorem it is necessary to prove that the conditional probabilities $P\{\xi_{t+1} = y|\xi_t = x\}$ or conditional probability rates $\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta}=y|\xi_t=x\}-P\{\xi_t=y|\xi_t=x\}}{\Delta}$ are equal to respectively $p(x, y)$ or $q(x, y)$ (or they are derivatives of this functions).

Theorem 8. Let $p : X_0 \times X \to \{f : \Upsilon \to [0, 1]\}$ be a discrete stochastic function. Let $(\xi_t)_{t\in\Upsilon}$ be a stochastic process defined by this stochastic function. Then for any $x \in X_0$ and $y \in X$ and $t \in \Upsilon$ we have: $P\{\xi_{t+1} = y|\xi_t = x\} = p(x, y)$

Proof. Let $(\Omega, \sigma(\Omega), P)$ be a probability space for the function p. Then, according to (5.5):

$$P\{\xi_{t+1} = y|\xi_t = x\} = \frac{P\{\varphi; \varphi(t) = x, \varphi(t + 1) = y\}}{P\{\varphi; \varphi(t) = x\}} = \tag{8.1}$$

$$= \frac{P^{x_0}_{[0,t)}(A^x_t) \cdot p(x, y)(t) \cdot P^y_{[t+1,T]}(\Omega^y_{[t+1,T]})}{P^{x_0}_{[0,t)}(A^x_t) \cdot \sum_{z\in\Omega} p(x, z)(t) \cdot P^y_{[t+1,T]}(\Omega^y_{[t+1,T]})} =$$

$$= \frac{p(x, y)(t)}{\sum_{z\in\Omega} p(x, z)(t)} = p(x, y)$$

□

Theorem 9. Let $q : X_0 \times X \to \{f : \Upsilon \to [0, 1]\}$ be a discrete intensity function. Let $(\xi_t)_{t\in\Upsilon}$ be a stochastic process defined by this stochastic function. Then for any $x \in X_0$ and $y \in X$ and $t \in \Upsilon$ we have:

$$\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = y|\xi_t = x\} - P\{\xi_t = y|\xi_t = x\}}{\Delta} = q(x, y)$$

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

Proof. Let $(\Omega, \sigma(\Omega), P)$ be a probability space for the function q. Let:

$$U_r(x,y)(t,t+\Delta) = \tag{8.2}$$

$$= \{\varphi \in \Omega; \varphi(t) = x, \varphi(t+\Delta) = y, \underset{t<t_1<t_2<\ldots<t_r=t+\Delta}{\exists} \varphi(t_i) \neq \varphi(t_{i+1})\}$$

In other words, $U_r(x,y)(t,t+\Delta)$ consists of these realizations in which x changes to y after exactly r significant transitions in time period $[t,t+\Delta)$. Using this notation we have:

$$\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = x | \xi_t = x\} - P\{\xi_t = x | \xi_t = x\}}{\Delta} = \tag{8.3}$$

$$= \lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = x | \xi_t = x\} - 1}{\Delta} =$$

$$= \lim_{\Delta\to 0} \frac{P(U_0(x,x)(t,t+\Delta)) - 1}{\Delta} + \sum_{r=2}^{0} \lim_{\Delta\to 0} \frac{P(U_r(x,x)(t,t+\Delta))}{\Delta}$$

According to (5.4) in definition 6 of probability P:

$$\lim_{\Delta\to 0} \frac{P(U_0(x,x)(t,t+\Delta)) - 1}{\Delta} = \lim_{\Delta\to 0} \frac{\exp\left(\int_t^{t+\Delta} q(x,x)(s)ds\right) - 1}{\Delta} = \tag{8.4}$$

$$\frac{d}{d\Delta} \exp\left(\int_t^{t+\Delta} q(x,x)(s)ds\right)\Big|_{\Delta=0} =$$

$$\exp\left(\int_t^t q(x,x)(s)ds\right) \cdot \frac{d}{d\Delta} \int_t^{t+\Delta} q(x,x)(s)ds\Big|_{\Delta=0} = q(x,x)(t)$$

For $x \neq y$:

$$\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = y | \xi_t = x\} - P\{\xi_t = y | \xi_t = x\}}{\Delta} = \tag{8.5}$$

$$\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = y | \xi_t = x\}}{\Delta} =$$

$$\lim_{\Delta\to 0} \frac{P(U_1(x,y)(t,t+\Delta))}{\Delta} + \lim_{\Delta\to 0} \sum_{r=2}^{0} \frac{P(U_1(x,y)(t,t+\Delta))}{\Delta}$$

If $y \in X_0$, then according to (5.5) in definition 6 of probability P:

$$\lim_{\Delta\to 0} \frac{P(U_1(x,y)(t,t+\Delta))}{\Delta} = \tag{8.6}$$

$$= \lim_{\Delta\to 0} \frac{\int_t^{t+\Delta} q(x,y)(s) \exp\left(\int_t^s q(x,x)(u)du + \int_s^{t+\Delta} q(y,y)(u)du\right)ds}{\Delta}\Big|_{\Delta=0} =$$

$$= q(x,y)(t) \exp\left(\int_t^t q(x,x)(u)du + \int_t^t q(y,y)(u)du\right) = q(x,y)$$

Similarly, for $x \neq y$ and $y \in X - X_0$ using equation (5.6):

$$\lim_{\Delta\to 0} \frac{P(U_1(x,y)(t,t+\Delta))}{\Delta} = \tag{8.7}$$

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

$$= \lim_{\Delta\to 0} \frac{\int_t^{t+\Delta} q(x,y)(s)\exp\Big(\int_t^s q(x,x)(u)du\Big)ds}{\Delta}\Big|_{\Delta=0} =$$

$$= q(x,y)(t)\exp\Big(\int_t^t q(x,x)(u)du\Big) = q(x,y)$$

For $r \geq 2$ it may be proved that:

$$P(U_r(x,y)(t,t+\Delta)) = \tag{8.8}$$

$$= \sum_{x\neq x, z\neq y} \int_t^{t+\Delta} q(x,z)(s)\exp\Big(\int_t^s q(x,x)(u)du\Big)P(U_{r-1}(z,y)(s,t+\Delta))ds$$

what allows a simple inductive proof that for all $r \geq 2$:

$$\lim_{\Delta\to 0} \frac{P(U_r(x,y)(t,t+\Delta))}{\Delta} = 0 \tag{8.9}$$

By (8.3), (8.4) and (8.9) we see that:

$$\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = x|\xi_t = x\} - P\{\xi_t = x|\xi_t = x\}}{\Delta} = q(x,x) \tag{8.10}$$

By (8.5), (8.6), (8.7) and (8.9) we see that:

$$\lim_{\Delta\to 0} \frac{P\{\xi_{t+\Delta} = y|\xi_t = x\} - P\{\xi_t = y|\xi_t = x\}}{\Delta} = q(x,y) \tag{8.11}$$

□

Using theorems similar to theorem 6 and 7, it can be proved that $P(\xi_{t+1} \in A|\xi_t = x) = \int_A p(x,y)dy$ for discrete-time, continuous-value stochastic process and $\lim_{\Delta\to 0} \frac{P\{(\xi_{t+\Delta}\in A-\{x\}|\xi_t=x)\}}{\Delta} = \int_A q(x,y)dy$.

9. Conclussions. The class of Markov stochastic processes defined by non-square stochastic and intensity functions is very wide. They may be used to model almost all biological phenomena. They are well-defined stochastic processes with well-defined probability spaces. These probability spaces are defined according to the biological definitions of the stochastic process, so they all have a clear interpretation. The probabilities are given by explicit formulas which allows for calculating probabilities for large numbers of realizations, albeit the calculations may sometimes prove to be complex or difficult. It is therefore possible to calculate many statistical characteristics of random variables defined on probability spaces as it is often done in biological applications.

Stochastic or intensity functions are not always calculated for biological models, although their Markov property is easy to show. Sometimes the calculations prove themselves to be too difficult, but in principle these functions always exists. For mathematicians, stochastic functions with non-square domains are not particularly elegant, perhaps because even the most common theorems do not translate into the non-square case. Chapman-Kolmogorov equation for Markov homogeneous processes is frequently presented as a product of stochastic matrices $P(t+s) = P(t)P(s)$ (where $P(v)$ is a collection of probabilities that state x transits to state y after v steps). It is clearly possible to formulate it this way only for square matrices. However, a more general version of this theorem may be formulated for the processes described in this article. In proof of the ergodic theorem, the square shape of the stochastic matrix

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

is a sufficient assumption. The consequences of the ergodic theorem are apparent in many biological models (right panel of Fig. 4) and one may say, that they are desirable. Parameters of the models are chosen so that a stabilization of the variables distribution during simulations is observed. But this stabilization is observed only for realizations which are not interrupted prior to the terminal time. Hence the processes defined by a stochastic function $p : X_0 \times X \to \{f : \Upsilon \to \mathbb{R}\}$ may can be analysed only using a stochastic processes induced by a function $p^{sq} : X_0 \times X_0 \to \{f : \Upsilon \to \mathbb{R}\}$ such that $p^{sq}(x, y)(t) = \frac{p(x,y)(t)}{\sum_{y \in X_0} p(x,y)(t)}$.

Kolmogorov equations are formed for such continuous-time discrete-value Markov stochastic processes, that functions $(t, s) \to P\{\xi_s = y|\xi_t = x\}$ are differentiable. Then function $\Phi_{x,y}(t, s) = P\{\xi_s = y|\xi_t = x\}$ with domain $\{(t, s); t < s\}$ satisfies the conditions:

$$\frac{\partial \Phi_{x,y}}{\partial s}(t, s) = \sum_z \Phi_{x,z}(t, s) q(z, y)(s) \tag{9.1}$$

$$\frac{\partial \Phi_{x,y}}{\partial t}(t, s) = -\sum_z q(x, z)(t) \Phi_{z,y}(t, s) \tag{9.2}$$

These equation may be shown to be true for non-square stochastic and intensity functions, but the theorems described in this article are more general, i.e. they do not require these functions to be differentiable. For infinite intensity matrices, the functions which are solutions to these equations may satisfy the condition $\sum_y \Phi_{x,y}(t, s) < 1$. This poses an obstacle in mathematical definition of a continuous-time discrete-value Markov stochastic processes, but in the biological case, the value $1 - \sum_y \Phi_{x,y}(t, s) < 1$ is a probability that the realizations of the stochastic process become infinite in a time period $[t, s)$. This is a result of having a broader space of events Ω, as well as an additional interpretation of an "explosive" Markov stochastic processes.

REFERENCES

[1] S. Asmussen, *Markov Chains. Applied Probability and Queues*, John Wiley and Sons Ltd., 2003.
[2] P. A. Ernst, L. D. Brown, L. Shepp, and R. L. Wolpert, *Stationary gaussian markov processes as limits of stationary autoregressive time series*, Journal of advances in Mathematics, 155 (2017), pp. 180–186.
[3] P. A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation*, John Wiley and Sons Ltd., 2017.
[4] G. C. Ghirardi, P. Pearle, and A. Rimini, *Markov processes in hilbert space and continuous spontaneous localization of systems of identical particles*, Physical Review A, 42 (1990), pp. 78–89.
[5] O. C. Ibe, *Markov Processes for Stochastic Modeling, 2nd ed.*, Chapman and Hall, 2013.
[6] M. Iosifescu, *Finite Markov processes and their applications*, John Wiley, 1980.
[7] M. E. Johnson, *Multivariate Statistical Simulation*, John Wiley and Sons, 1987.
[8] N. J. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions, Volume 1, 2nd Ed.*, John Wiley, 1995.
[9] J. Kelemen and P. Sosík, *Advances in Artificial Life. 6th European Conference*, ECAL, Prague, Czech Republic, 2001.
[10] A. K. Klenke, *Probability Theory: A Comprehensive Course*, Springer, 2013.
[11] M. Komosinski and A. Adamacky, *Artificial Life Models in Software*, Springer, 2005.
[12] K. Li, X. Li, G. W. Irwin, and G. He, *Life System Modelling and Simulation, International Conference*, LSMS, Shanghai, China, 2007.
[13] J. R. Norris, *Markov Chains*, Cambridge University Press, 1998.

ISSN: 2347-1921
Volume: 14 Issue: 01
Journal of Advances in Mathematics

[14] A. E. Raftery, *A model for high-order markov chains*, Journal of the Royal Statistical Society. Series B (Methodological), 47 (1985), pp. 528–539.
[15] A. Rønn-Nielsen and E. Hansen, *Conditioning and Markov properties*, Department of Mathematical Sciences, University of Copenhagen, on line: http://www.math.ku.dk/noter/filer/beting.pdf, 2014.
[16] M. Sokol, *A probability space of continuous-time discrete value stochastic process with markov property*, Journal of advances in Mathematics, 12 (2016), pp. 5975–5990.
[17] S. Tuljapurkar, *Stochastic matrix models*, in Population and Community Biology Series 18, M. B. Usher, D. L. DeAngelis, and B. J. F. Manly, eds., Springer Nature, 2007.