UDC 519.2

Markov stochastic processes in biology – almost the same than in mathematics but a bit different

M. Sokół*

* *Faculty of Biology, Biological and Chemical Research Centre, University of Warsaw, Zwirki i Wigury 101, 02-089 Warsaw, Poland*

Abstract. Due to annual and diurnal fluctuations almost all stochastic processes in biology are non-homogeneous. They can be defined by any function $p : W \times X \to \mathcal{F}$ or $q : W \times X \to \mathcal{F}$ where X is a set of states, W is a limited subset of X, $\mathcal{F}$ is a set of time-depended integrable functions $f : [0, T) \to [0, \infty)$. At any time t for any state $w \in W$ the sum $\sum_X p(w, x)(t)$ or the integration $\int_X p(w, x)(t)dx$ are equal to 1 (discrete time processes), and the sum $q(w, x)(t)$ is equal 0 or the integration $\int_X q(w, x)(t)dx$ (continuous time processes) is positive but finite. This is not only obvious generalization of a stochastic/intensity matrix/field but also modification of them. Limitation of W is significant but X including all possible states is often infinite. These Markov stochastic processes have non-quadratic stochastic matrices and fields. It is possible because realizations of these processes are interrupted when the state after transitions is outside of W.

For all function $p : W \times X \to \mathcal{F}$ and $q : W \times X \to \mathcal{F}$ there exist stochastic processes with Markov property. It can be simulated by computer programs. It has correctly defined probability space while for some infinite quadratic matrices it is not true.

Keywords: stochastic matrix, stochastic field, intensity matrix, intensity field.

1. Introduction

A stochastic matrix is a square matrix, which entries are non-negative real numbers and each row (right matrix) or column (left matrix) summing to 1 [2]. Intensity matrixes is a square matrix, which entries are non-negative real numbers, expect diagonal with non-positive entries. Each row (right matrix) or column (left matrix) of this matrix summing to 0 [5]. Their continuous equivalents are named random fields [1], [3]. These lexical definitions of stochastic and intensity matrices and fields are insufficient for the applications.

Due to diurnal or seasonal fluctuations of a probability it is necessary to consider the non-homogeneous Markov processes. The entries in stochastical and intensity matrices must be replaced by time-depended functions, which values are probabilities or probability rates. The set of such function will be noted as $\mathcal{F}$.

2. Modification

The biologists calculate probabilities inserted in stochastic or intensity matrix using a regression between value of state and frequency of events. General rule keeping in biology runs: "it is not possible to prolong the regressions beyond the observed range". Biological systems are too complicated to predict what it is happening for extremal values of state. It is seen for a simple birth and death process in discrete time applied to biological populations (figure 1). The same rules apply to stochastic fields,

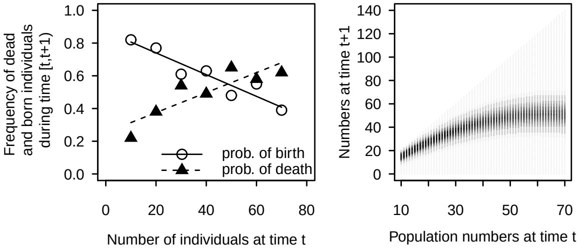

Figure 1. On left - the regressions between population numbers k at time t and frequency of born and dead individuals during time $[t, t+1)$ calculated on interval $[10, 70]$ only. For $k < 10$ and $k > 70$ the value of probability of birth and death are impossible to foresee. The formulas $p_r(k)$ and $p_s(k)$ allow calculating a probability of change population numbers 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 shown on right. The darker points indicate greater value of probabilities. They have positive value for $k \in [10, 70]$ and $n \in [0, 2k]$. Minimal size of this matrix is $[10, 70] \times [0, 140]$.

intensity matrices and intensity fields. The set of states, for which probabilities or probability rates can be estimated is limited. It will be named W. The set for finish states cannot be limited.

3. Definitions of stochastic and intensity "matrix"/"field"

The terms "matrix" and "field" suggest two-dimensional objects. But, when a state is a vector $w = (n_1, n_2, ...)$, then the ordering of the states is not natural. For discrete-time discrete-value processes it is better to consider a function $p : W \times X \to \mathcal{F}$, where X is a set of events, $W \subset X$ is limited and $p(w,v)(t) \geq 0$ for all $w \in W$ and $v \in X$, and:

$$\sum_{v \in X} p(w,v)(t) = 1 \quad \text{for all } w \in W \text{ and } t \in [0, T).$$

For continuous-time discrete-value processes an intensity matrix may be changed to a function $q : W \times X \to \mathcal{F}$, such that $q(w, v)(t) \geq 0$ if $w \neq v$ and $q(w, w)(t) \leq 0$, and:

$$\sum_{v \in X} q(w, v)(t) = 0 \quad \text{for all } w \in W \text{ and } t \in [0, T).$$

Functional view on described terms causes, that a simple stochastic matrix would be write in another orientation than we has usually done it (figure 1, on right).

For discrete-time continuous-value processes a stochastic field is a function $p : W \times X \to \mathcal{F}$, such that $p(w, v)(t) \geq 0$ for all $w \in W$ and $v \in X$. and:

$$\int_X p(w, v)(t) dv = 1 \quad \text{for all } w \in W \text{ and } t \in [0, T).$$

For continuous-time continuous-value processes an intensity field is a function $q : W \times X \to \mathcal{F}$, such that $q(w, v)(t) \geq 0$ for all $w \in W$ and $v \in X$ and there exist such $0 < M < \infty$ that:

$$\int_X q(w, v)(t) dv < M \quad \text{for all } w \in W \text{ and } t \in [0, T).$$

4. Simulation and probability space

A simulation of a Markov processes for stochastic matrix is shown on figure 2. There exist simulators of Markov stochastic processes for each stochastic intensity matrix fields.

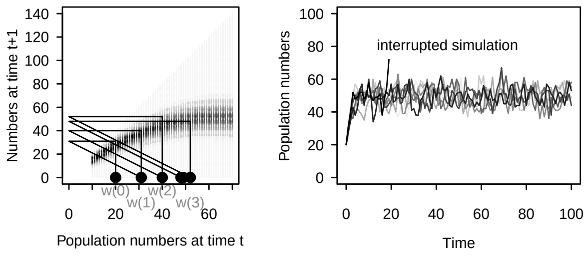

Figure 2. On left - the stochastic matrix for birth and death process in discrete time. On right - the sample of the eight realizations of this process to maximal time $T = 100$. When the population receives numbers greater than 70 the simulation is interrupted. This phenomena is interpreted as the population input into an unknown system of regulatory mechanisms.

By a repeating of the simulation a set Ω of all possible realization is formed. There exists a probability space $(\Omega, \sigma(\Omega), P)$ such that for all $A \in \sigma(\Omega)$ the $P(A)$ is a probability that random realization including to A. Construction of this probability space is similar to Lebesgue measure definition. It has already published for an intensity matrix [4]. It formes a probability space for the stochastic process $(\xi_t)_{t\in[0,T]}$ such that $\xi_t : \Omega \ni \varphi \to \varphi(t) \in X$.

References

1. *Adler R. J., Taylor J. E.* Random Fields and Geometry. — Springer Science and Business Media, 2009.
2. *Asmussen S. R.* Applied Probability and Queues. — Application of Mathematics, Stochastic Modelling and Applied Probability. — Vol. 51. — Springer. 2003.
3. *Khoshnevisan D.* Multiparameter Processes: An Introduction to Random Fields. — Springer Science and Business Media, 2006.
4. *Sokół M.* A probability space of continuous-time discrete value stochastic process with Markov property // Journal of Advancis in Mathematics. — 2016. — Vol.12 — P. 5975–5991.
5. *Syski R.* Passage Times for Markov Chains. — IOS Press, 1992.