

МОДЕЛЬ РОСТА ОДНОТИПНОЙ ПОПУЛЯЦИИ В ПРОСТРАНСТВЕ И ВРЕМЕНИ

Бутов А.А., Егоров А.Г.

ФГБОУ ВПО «Ульяновский государственный университет», Ульяновск, Россия (432017, Российская Федерация, город Ульяновск, улица Льва Толстого, дом 42), e-mail: al.g.egorov@gmail.com

При исследовании популяций, развивающихся в пространстве и времени, возникает задача оценки численности популяции в произвольных областях, зная начальное распределение. Под однотипной популяцией в данной работе рассматривается популяция бактерий. Предполагается, что наблюдается не вся область, а лишь выборочные ее сектора в случайные моменты времени. При этом наблюдение сильно зашумлено – в момент подсчета числа бактерий в некоторой области бактерии могут перемещаться в соседние области. Целью данной работы является построение адекватной математической модели, позволяющей анализировать развитие однотипной популяции в пространстве и времени. Методом исследования является имитационное стохастическое моделирование. Оценка реального числа бактерий в секторе по серии наблюдений осуществляется на основе дискретной схемы Калмана.

Ключевые слова: случайное блуждание, процессы рождения и гибели, дискретная схема Калмана.

THE MODEL OF GROWTH ONE-TYPE POPULATION IN SPACE AND TIME

Butov A.A., Egorov A.G.

Ulyanovsk State University, Ulyanovsk, Russia (432970, Ulyanovsk, street L. Tolstoy, 42), e-mail: al.g.egorov@gmail.com

While researching the developing of a population in space and time the task of population size evaluation in any area based on the original distribution. In this paper we consider of a bacteria population. We supposed that under observation is not all the plane, but some areas and it happens in random time moments. Wherein the observations are noisy, associated with movement of bacteria in the other areas. The aim of this work is to create an appropriate mathematical model allowing to make analysis of one-type population evolution in space and time. The method of researches is the imitation of stochastic modeling. The estimation of real bacteria number in the selected area is the result of observation series based on discrete Kalman filter.

Keywords: random walk, birth-and-death processes, discrete Kalman filter.

Введение

При исследовании популяций, развивающихся в пространстве и времени, возникает задача оценки численности популяции в произвольных областях в некоторый момент времени, зная начальное распределение [1]. Под однотипной популяцией в данной работе рассматривается популяция бактерий. Предполагается, что наблюдается не вся область, а лишь выборочные ее сектора в случайные моменты времени. При этом наблюдение сильно зашумлено – в момент подсчета числа бактерий в некоторой области бактерии могут перемещаться в соседние области. По этой причине разработка алгоритмов и математических моделей, предназначенных для анализа роста однотипных популяций, представляет актуальную научную проблему. Целью данной работы является построение адекватной математической модели, позволяющей анализировать развитие однотипной популяции в пространстве и времени. Методом исследования является имитационное стохастическое моделирование.

Оценка реального числа бактерий в секторе на основе наблюдений осуществляется с помощью дискретной схемы Калмана.

Описание модели

Рассмотрим математическую модель роста одноклеточной популяции, развивающуюся из исходной популяции во времени и двумерном пространстве. Для определенности опишем процесс, являющийся моделью роста популяций бактерий.

Предположим, что ареал Λ имеет форму прямоугольника. Разобьем ареал на одинаковые прямоугольные области $S^{i,j}$, $1 \leq i \leq L$ и $1 \leq j \leq M$, L и M – натуральные числа, таким образом $\Lambda = \{S^{i,j}\}$, $1 \leq i \leq L$ и $1 \leq j \leq M$.

В каждой области $S^{i,j}$, $1 \leq i \leq L$ и $1 \leq j \leq M$ может находиться неотрицательное целое число бактерий. Каждая бактерия может с вероятностью ρ ($0 \leq \rho \leq 1$) поделиться на две. Обозначим $B_t^{i,j}$ процесс, считающий количество появившихся бактерий в результате деления внутри области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$). Предполагается, что интенсивность деления пропорциональна количеству бактерий в области в момент времени t с коэффициентом пропорциональности ρ , т.е. компенсатор процесса $B_t^{i,j}$ описывается классической моделью Мальтуса [3] и равен:

$$\tilde{B}_t^{i,j} = \int_0^t \rho \cdot N_s^{i,j} ds \quad (1)$$

где $N_t^{i,j}$ – количество бактерий в момент времени t внутри области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$). Основные определения и термины для описания семи мартингалов см. в [2].

Рассмотрим также процессы перехода бактерий в область $S^{i,j}$ из соседних областей и ухода бактерий из области $S^{i,j}$ в соседние области. Обозначим их $A_t^{i,j}$ и $D_t^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$) соответственно. Предполагается, что компенсатор процесса ухода $D_t^{i,j}$ бактерии из области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$) также пропорционален количеству бактерий в области в момент времени t с коэффициентом пропорциональности δ .

$$\tilde{D}_t^{i,j} = \int_0^t \delta \cdot N_s^{i,j} ds, \quad (2)$$

где δ – интенсивность ухода бактерии из области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$).

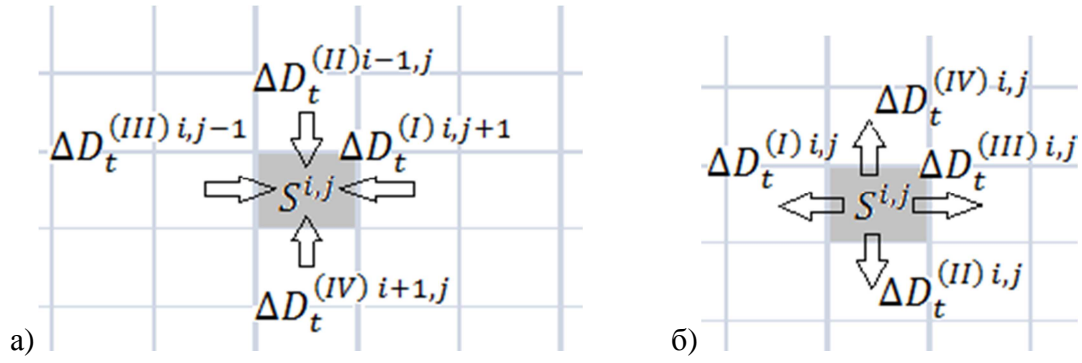


Рисунок 1. Переходы бактерий из области $S^{i,j}$ и в область $S^{i,j}$ из соседних областей

Обозначим $\Delta D_t^{(I) i,j}$ – количество бактерий, перешедших из области $S^{i,j}$ в область $S^{i,j-1}$, $1 \leq i \leq L$ и $1 < j \leq M$ в момент времени t ; $\Delta D_t^{(II) i,j}$ – из области $S^{i,j}$ в область $S^{i+1,j}$, $1 \leq i < L$ и $1 \leq j \leq M$; $\Delta D_t^{(III) i,j}$ – из области $S^{i,j}$ в область $S^{i,j+1}$, $1 \leq i \leq L$ и $1 \leq j < M$ и $\Delta D_t^{(IV) i,j}$ – из области $S^{i,j}$ в область $S^{i+1,j}$, $1 \leq i < L$ и $1 \leq j \leq M$, см. рисунок 1 (а, б). Пусть бактерии не умирают, тогда выполняется соотношение:

$$\Delta D_t^{i,j} = \Delta D_t^{(I) i,j} + \Delta D_t^{(II) i,j} + \Delta D_t^{(III) i,j} + \Delta D_t^{(IV) i,j} \quad (3)$$

Предполагается, что бактерия не может одновременно перейти в две разных области, т. е. выполняется равенство (4).

$$\sum_{\substack{l,s=I, \\ l \neq s}}^{IV} \Delta D_t^{(s) i,j} \cdot \Delta D_t^{(l) i,j} = 0 \quad \text{P-п. н.} \quad (4)$$

Предположим, что переходы бактерий из области $S^{i,j}$ будут происходить в момент скачков пуассоновского процесса $\pi_t^{i,j}$. Обозначим эти моменты $\tau_k^{i,j}$, $k > 0, 1 \leq i < L$ и $1 \leq j \leq M$. Тогда вероятности переходов будут равны:

$$\left\{ \begin{array}{l} P\left(\Delta D_{\tau_k^{i,j}}^{(I) i,j} \middle| \mathcal{F}_{\tau_k^{i,j}-}\right) = \alpha^{i,j}, \\ P\left(\Delta D_{\tau_k^{i,j}}^{(II) i,j} \middle| \mathcal{F}_{\tau_k^{i,j}-}\right) = \beta^{i,j}, \\ P\left(\Delta D_{\tau_k^{i,j}}^{(III) i,j} \middle| \mathcal{F}_{\tau_k^{i,j}-}\right) = \kappa^{i,j}, \\ P\left(\Delta D_{\tau_k^{i,j}}^{(IV) i,j} \middle| \mathcal{F}_{\tau_k^{i,j}-}\right) = \varepsilon^{i,j}, \\ \alpha^{i,j} + \beta^{i,j} + \kappa^{i,j} + \varepsilon^{i,j} = 1 \end{array} \right. \quad (5)$$

С учетом (5) процесс $A_t^{i,j}$ равен

$$A_t^{i,j} = \int_0^t (\Delta D_s^{i,j-1} + \Delta D_s^{i+1,j} + \Delta D_s^{i,j+1} + \Delta D_s^{i-1,j}) \quad (6)$$

Предположим, что ареал замкнут всюду, т. е. $\alpha^{i,1} = 0 \forall i, 1 \leq i \leq L$; $\beta^{L,j} = 0 \forall j, 1 \leq j \leq M$; $\kappa^{i,M} = 0 \forall i, 1 \leq i \leq L$ и $\varepsilon^{1,j} = 0 \forall j, 1 \leq j \leq M$.

С учетом формул (1-3, 6) получаем, что количество бактерий в области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$) в момент времени t равно

$$N_t^{i,j} = N_0^{i,j} + B_t^{i,j} + A_t^{i,j} - D_t^{i,j} \quad (7)$$

где $N_0^{i,j}$ – начальное количество бактерий в из области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$).

Процесс наблюдения $Q_t^{i,j}$ за процессом $N_t^{i,j}$ будет выглядеть следующим образом

$$\begin{aligned} Q_t^{i,j} &= Q_0^{i,j} + \int_0^t (N_s^{i,j} - Q_{s-}^{i,j} + \xi_s) d\pi_s, \forall (i,j) \in U \\ Q_0^{i,j} &= N_0^{i,j}, \forall (i,j) \in U, \end{aligned} \quad (8)$$

где ξ_t – целочисленная случайная величина, распределенная по закону Пуассона с параметром $\mu > 0$, ошибка измерений.

Предположим, что распределение бактерий в начальный момент времени по областям ареала Λ неравномерно. Требуется по наблюдению $Q_t^{i,j}$ оценить численность бактерий в каждой области ареала.

Теоретическая оценка математического ожидания числа бактерий

Математическое ожидание процесса $N_t^{i,j}$, считающего количество бактерий в области $S^{i,j}$ ($1 \leq i \leq L$ и $1 \leq j \leq M$), из уравнений (1, 2, 6, 7) равно:

$$\begin{aligned} EN_t^{i,j} &= N_0^{i,j} + \int_0^t \rho \cdot EN_s^{i,j} ds + \\ &+ E \int_0^t (\Delta D_s^{i,j-1} + \Delta D_s^{i+1,j} + \Delta D_s^{i,j+1} + \Delta D_s^{i-1,j}) - \int_0^t \delta \cdot EN_s^{i,j} ds = \\ &= N_0^{i,j} + \int_0^t \rho \cdot EN_s^{i,j} ds + \alpha^{i,j+1} \cdot \int_0^t \delta \cdot EN_s^{i,j+1} ds + \beta^{i-1,j} \cdot \int_0^t \delta \cdot EN_s^{i-1,j} ds + \\ &+ \kappa^{i,j-1} \cdot \int_0^t \delta \cdot EN_s^{i,j-1} ds + \varepsilon^{i+1,j} \cdot \int_0^t \delta \cdot EN_s^{i+1,j} ds - \int_0^t \delta \cdot EN_s^{i,j} ds, \quad (9) \end{aligned}$$

где $1 \leq i \leq L$ и $1 \leq j \leq M$, причем слагаемые с индексами: 0 , $L + 1$ и $M + 1$ равны 0 .
Запишем полученные уравнения (9) в дифференциальном виде:

$$\begin{aligned} \frac{d(EN_t^{i,j})}{dt} = & (\rho - \delta) \cdot EN_t^{i,j} + \alpha^{i,j+1} \cdot \delta \cdot EN_t^{i,j+1} + \beta^{i-1,j} \cdot \delta \cdot EN_t^{i-1,j} + \\ & + \kappa^{i,j-1} \cdot \delta \cdot EN_t^{i,j-1} + \varepsilon^{i+1,j} \cdot \delta \cdot EN_t^{i+1,j} \end{aligned} \quad (10)$$

Получена система $L \times M$ однородных линейных дифференциальных уравнений (10) с $L \times M$ неизвестными и начальными значениями $N_0^{i,j}$. Решение таких систем подробно изложено в [4]. Решение в общем виде можно записать через матричную экспоненту:

$$EN_t = N_0 \cdot e^{tA} \quad (11)$$

где N_t – вектор, $N_t = (N_t^{1,1}, N_t^{1,2}, \dots, N_t^{1,M}, N_t^{2,1}, N_t^{2,2}, \dots, N_t^{2,M}, \dots, N_t^{L,1}, N_t^{L,2}, \dots, N_t^{L,M})$; A – матрица, составленная из коэффициентов уравнений (10); N_0 – начальное значение вектора N_t .

Оценка численности числа бактерий

Предположим, что наблюдение за областями $S^{i,j}$ $1 \leq i \leq L$ и $1 \leq j \leq M$ происходит в случайные моменты времени t_k , $k \geq 0$ скачков пуассоновского процесса π_t с интенсивностью $\lambda > 0$ одновременно во всех областях $S^{i,j}$. И предположим также, что при подсчете количества бактерий в области $S^{i,j}$ $1 \leq i \leq L$ и $1 \leq j \leq M$ происходит ошибка $\eta_k^{i,j}$, т. е.

$$Q_{t_k}^{i,j} = N_{t_k}^{i,j} + \eta_k^{i,j} \quad (11)$$

где $\eta_k^{i,j}$, $k \geq 0$ – совокупность независимых случайных величин, с нулевым средним и дисперсией σ_η^2 .

Значение процесса $N_t^{i,j}$ в моменты времени t_k будет отличаться от математического ожидания $EN_t^{i,j}$ на некоторую мартингальную часть. В предыдущем пункте был рассмотрен способ, позволяющий отыскать теоретическую оценку математического ожидания числа бактерий. Само решение зависит от конкретных значений интенсивностей переходов ρ и δ и весовых коэффициентов $\alpha^{i,j}$, $\beta^{i,j}$, $\kappa^{i,j}$, $\varepsilon^{i,j}$, $1 \leq i \leq L$, $1 \leq j \leq M$, поэтому обозначим, что от момента t_k до момента t_{k+1} математическое ожидание числа бактерий меняется по закону:

$$EN_{t_{k+1}}^{i,j} = EN_{t_k}^{i,j} + \Delta u_{t_k} \quad (12)$$

где Δu_{t_k} – некоторая функция, которая ищется исходя из решений уравнений (10). Суммируя вышесказанное, получаем формулу (13):

$$N_{t_{k+1}}^{i,j} = N_{t_k}^{i,j} + \Delta u_{t_k} + \xi_{k+1} \quad (13)$$

где ξ_k – независимые одинаково распределенные случайные величины с нулевым средним и дисперсией σ_ξ^2 .

Рассмотрим способ оценки реального числа бактерий $\hat{Q}_{t_k}^{i,j}$ по наблюдению за областью $S^{i,j}$ $1 \leq i \leq L$ и $1 \leq j \leq M$ на основе дискретной схемы Калмана [5].

В каждый момент времени оценка реального числа бактерий в области $S^{i,j}$ будет считаться как средневзвешенное между теоретической оценкой бактерий и наблюдаемыми значениями с коэффициентами Z_{k+1} и $(1 - Z_{k+1})$ соответственно, где Z_k – коэффициенты Калмана [5].

$$\begin{cases} \hat{Q}_{t_{k+1}}^{i,j} = Z_{k+1} \cdot Q_{t_k}^{i,j} + (1 - Z_{k+1}) \cdot (\hat{Q}_{t_k}^{i,j} + u_k) \\ \hat{Q}_{t_0}^{i,j} = Q_{t_0}^{i,j} \end{cases} \quad (14)$$

где Z_{k+1} – коэффициент Калмана, $Z_{k+1} \in [0,1]$. Нужно так подобрать коэффициенты Z_{k+1} , чтобы получившаяся оценка числа бактерий $\hat{Q}_{t_{k+1}}^{i,j}$ была наиболее близкой к реальному числу $N_{t_{k+1}}^{i,j}$. К примеру, если известно, что ошибка обнаружения при подсчете числа бактерий очень мала, то коэффициенту Z_{k+1} нужно придать больший вес. И наоборот, если известно, что значение числа бактерий при подсчете очень не точное, следует больше ориентироваться на теоретически предсказанное значение $\hat{Q}_{t_k}^{i,j} + u_k$.

Если обозначить ошибку $e_{t_{k+1}}^{i,j}$ между реальным значением числа бактерий $N_{t_{k+1}}^{i,j}$ в области $S^{i,j}$ и их оценкой $\hat{Q}_{t_{k+1}}^{i,j}$, тогда задачу оценивания можно переформулировать в задачу минимизации среднеквадратической ошибки в момент времени t_{k+1} , см. формулу (16).

$$e_{t_{k+1}}^{i,j} = N_{t_{k+1}}^{i,j} - \hat{Q}_{t_{k+1}}^{i,j} \quad (15)$$

$$E(e_{t_{k+1}}^{i,j}{}^2) \rightarrow \min \quad (16)$$

Теорема. При сделанных выше предположениях оценка числа бактерий $\hat{Q}_{t_{k+1}}^{i,j}$ будет оптимальной при:

$$Z_{k+1} = \frac{E(e_{t_k}^{i,j^2}) + \sigma_\xi^2}{E(e_{t_k}^{i,j^2}) + \sigma_\xi^2 + \sigma_\eta^2}$$

Доказательство

Используя уравнения (11), (13) и (14), перепишем уравнение (15) в виде:

$$\begin{aligned} e_{t_{k+1}}^{i,j} &= N_{t_{k+1}}^{i,j} - \hat{Q}_{t_{k+1}}^{i,j} = N_{t_{k+1}}^{i,j} - Z_{k+1} \cdot Q_{t_{k+1}}^{i,j} - (1 - Z_{k+1}) \cdot (\hat{Q}_{t_k}^{i,j} + u_k) = \\ &= N_{t_k}^{i,j} + u_k + \xi_{k+1} - Z_{k+1} \cdot (N_{t_k}^{i,j} + u_k + \xi_{k+1} + \eta_{k+1}) - (1 - Z_{k+1}) \cdot (\hat{Q}_{t_k}^{i,j} + u_k) = \\ &= (1 - Z_{k+1}) \cdot (N_{t_k}^{i,j} - \hat{Q}_{t_k}^{i,j} + \xi_{k+1}) - Z_{k+1} \cdot \eta_{k+1} = \\ &= (1 - Z_{k+1}) \cdot (e_{t_k}^{i,j} + \xi_{k+1}) - Z_{k+1} \cdot \eta_{k+1} \end{aligned}$$

Из предположения независимости случайных величин, входящих в $e_{t_{k+1}}^{i,j}$, и с учетом того, что $E(\xi_k) = 0$ и $E(\eta_k) = 0$, получаем, что выражение (15) можно переписать в виде:

$$E(e_{t_{k+1}}^{i,j^2}) = (1 - Z_{k+1})^2 \cdot (E(e_{t_k}^{i,j^2}) + \sigma_\xi^2) + Z_{k+1}^2 \cdot \sigma_\eta^2$$

Это выражение принимает минимальное значение при

$$Z_{k+1} = \frac{E(e_{t_k}^{i,j^2}) + \sigma_\xi^2}{E(e_{t_k}^{i,j^2}) + \sigma_\xi^2 + \sigma_\eta^2} \quad (17)$$

Среднеквадратичная ошибка, минимизирующая значения коэффициентов Калмана, при этом равна:

$$\begin{cases} E(e_{t_{k+1}}^{i,j^2}) = \sigma_\eta^2 \cdot \frac{E(e_{t_k}^{i,j^2}) + \sigma_\xi^2}{E(e_{t_k}^{i,j^2}) + \sigma_\xi^2 + \sigma_\eta^2} \\ E(e_{t_0}^{i,j^2}) = E(\eta_0^2) = \sigma_\eta^2 \end{cases} \quad (18)$$

Что и требовалось доказать.

Заключение

Рассмотренная в статье модель позволяет описать рост однотипной популяции в пространстве и времени на основе семимартингалных точечных процессов по известному начальному распределению. Получена итерационная формула (17) для вычисления оптимальной оценки числа бактерий в произвольной области на основе нескольких наблюдений.

Работа выполнена при поддержке Министерства образования и науки РФ (НИР, проводимые в рамках государственного задания Министерства образования и науки РФ на 2013 г., Программы стратегического развития УлГУ на 2012–2016 гг., Программы развития деятельности студенческих объединений УлГУ на 2012-2013 гг.).

Список литературы

1. Арзамасцев А. А., Соломина О. А., Моделирование роста биологической популяции на плоскости // Математическое моделирование. – 2009. – Т. 21, № 4.
2. Липцер Р. Ш., Ширяев А. Н. Статистика случайных процессов. – М.: Наука, 1974. – 696 с.
3. Разжевайкин В. Н. Модели динамики популяций // Вычислительный центр им. А.А. Дородницына, Российской академии наук, 2006.
4. Романко В.К. Курс дифференциальных уравнений и вариационного исчисления. – 2-е изд. – М.: Лаборатория Базовых Знаний, 2001. – 344 с.
5. Kalman R. E. A new approach to linear filtering and prediction problems // Trans. ASME series D.,J. Basic Engg. 1960. V. 82. P. 35-45.

Рецензенты:

Кемер А.Р., д.ф.-м.н., профессор, ФГБОУ ВПО «Ульяновский государственный университет», г. Ульяновск.

Андреев А.С., д.ф.-м.н., профессор, ФГБОУ ВПО «Ульяновский государственный университет», г. Ульяновск.