

ДИСКРЕТНЫЕ СТОХАСТИЧЕСКИЕ МОДЕЛИ
ДЛЯ ИССЛЕДОВАНИЯ
ДИНАМИКИ РАСПРОСТРАНЕНИЯ
ИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ

Перцев Н.В., Леоненко В.Н.

Омский филиал Института математики
им. С.Л.Соболева СО РАН

Работа поддержана РФФИ (проект № 09-01-00098-а) и СО РАН (проект № 26)

В докладе представлена методология построения стохастических моделей социально значимых инфекционных заболеваний на основе стохастических рекуррентных уравнений.

К таким заболеваниям относятся туберкулез, СПИД и некоторые другие заболевания.

В своей работе мы исходили из имеющихся моделей, разработанных сотрудниками ИВМ РАН в соавторстве со специалистами из других институтов и организаций:

- 1. Perelman M.I., Marchuk G.I., Borisov S.E., et. al. Tuberculosis epidemiology in Russia: the mathematical model and data analysis // Russ. J. Numer. Anal. Math. Modelling. 2004, V.19. N.4. P.305–314.**
- 2. Авилов К.К., Романюха А.А. Математические модели распространения и контроля туберкулёза (обзор) // Математическая биология и биоинформатика. 2007. Т.2. N.2. С.188–318.**
- 3. Nosova E.A., Romanyukha A.A. Regional index of HIV infection Risk based on factors of social disadaptation // Russ. J. Numer. Anal. Math. Modelling. 2009. V.24. N.4. P.325–340.**
- 4. Романюха А.А., Носова Е.А. Модель распространения ВИЧ–инфекции в результате социальной дезадаптации // Управление большими системами. 2011. Вып.34. С.227–253.**

Обширный класс стохастических моделей опирается на марковские случайные процессы с дискретным временем.

Для этих моделей основными проблемами являются:

- формализация состояний изучаемого процесса;
- задание переходных вероятностей P_{ij} из состояния i в состояние j за один шаг времени;
- аналитическое исследование;
- разработка моделирующих программ;
- подбор параметров моделей по реальным данным;
- проведение вычислительных экспериментов, направленных на исследование конкретных содержательных задач.

В основе нашего подхода лежат четыре основных блока.

БЛОК 1. Используется дискретное время с шагом 1 сутки, что соответствует жизненному циклу человека, например, ежедневные контакты, возможна неоднородность по времени (будние дни, выходные дни, сезонность времени года).

БЛОК 2. Вместо задания переходных вероятностей P_{ij} применяются стохастические рекуррентные уравнения, которые связывают между собой численности различных групп населения,

стохастическая модель может быть построена на основе аналога в виде уже существующей системы дифференциальных уравнений,

переменные модели целочисленны,

возможно высоко-размерное и многопараметрическое задание различных групп населения,

возможно детальное моделирование отдельных стадий, которые проходит индивидум в процессе развития и лечения болезни («модель в модели»).

БЛОК 3. Построение вспомогательных детерминированных систем рекуррентных уравнений, позволяющих оценивать динамику математических ожиданий численностей изучаемых групп населения.

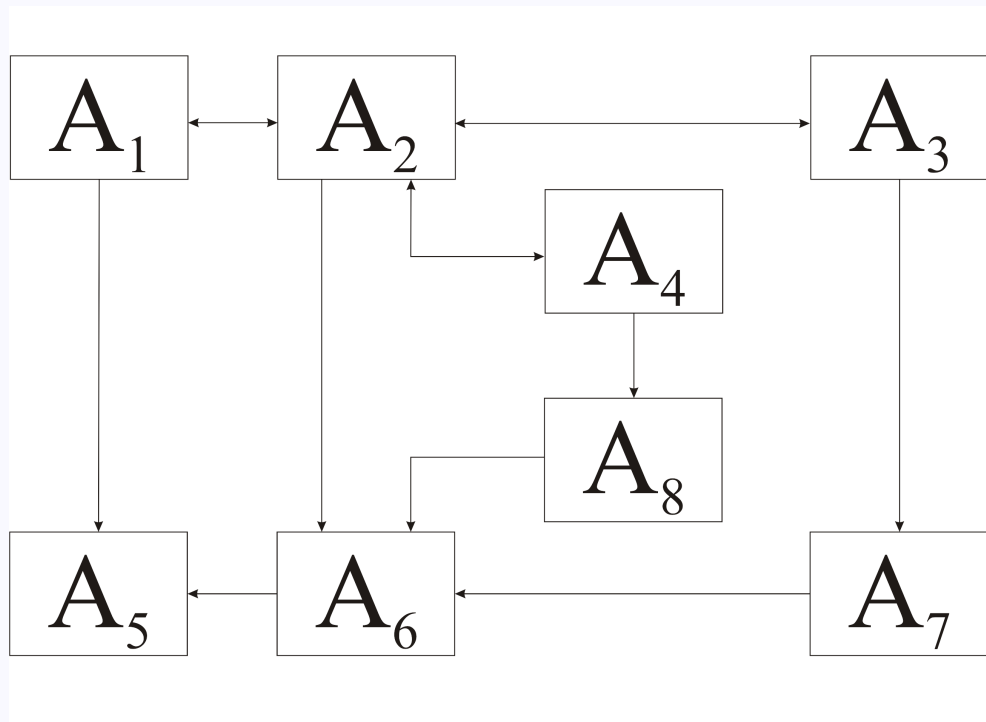
БЛОК 4. Вычисления проводятся на основе генерации совокупности случайных величин с биномиальным и мультиномиальным законами распределения с помощью эффективных алгоритмов статистического моделирования и хорошо проверенных датчиков псевдослучайных чисел.

Предложенный подход ранее был использован для исследования динамики распространения туберкулеза в регионах России:

1. Перцев Н.В., Романюха А.А., Касаткина В.С. Нелинейная стохастическая модель распространения туберкулеза // Системы управления и информационные технологии. 2008. N.1.2(31). С.246–250.
2. Pertsev N.V., Leonenko V.N. Stochastic individual-based model of spread of tuberculosis // Russ. J. Numer. Anal. Math. Modelling. 2009. V.24. N.4. P.341–360.
3. Перцев Н.В., Пичугин Б.Ю. Индивидуум-ориентированная стохастическая модель распространения туберкулеза // Сибирский журнал индустриальной математики. 2009. Т.12. N.2(38). С.97–110.

Ниже приводится стохастическая модель распространения ВИЧ, которая основана на модели в форме системы нелинейных дифференциальных уравнений, предложенной А.А.Романюхой и Е.А.Носовой. В этой модели используется описание структуры населения регионов России с точки зрения степени социальной дезадаптации индивидуумов.

Схема модели (не показаны притоки индивидуумов и оттоки индивидуумов вследствие смертности и миграции):



Переменные модели:

$x_1(t)$ – численность социально адаптированных индивидуумов, восприимчивых к ВИЧ (группа A_1) в м. вр. t

$x_2(t)$ – численность восприимчивых к ВИЧ индивидуумов с высоким риском развития патологии (группа A_2) в м. вр. t

$x_3(t)$ – численность восприимчивых к ВИЧ индивидуумов с установленным хроническим алкоголизмом (группа A_3) в м. вр. t

$x_4(t)$ – численность восприимчивых к ВИЧ индивидуумов с установленной наркотической зависимостью (группа A_4) в м. вр. t

$x_5(t)$ – численность социально адаптированных индивидуумов, инфицированных ВИЧ (группа A_5) в м. вр. t

$x_6(t)$ – численность индивидуумов, инфицированных ВИЧ, с высоким риском развития патологии (группа A_6) в м. вр. t

$x_7(t)$ – численность индивидуумов, инфицированных ВИЧ, с установленным хроническим алкоголизмом (группа A_7) в м. вр. t

$x_8(t)$ – численность индивидуумов, инфицированных ВИЧ, с установленной наркотической зависимостью (группа A_8) в м. вр. t

Уравнение для $x_1(t)$, (t и все $x_i(t - 1)$ – фиксированы):

$$x_1(t) = \hat{x}_1(t) - u_{12}(t) - u_{15}(t) + u_{21}(t) + f_1(t),$$

$\hat{x}_1(t) \sim Bin(x_1(t - 1), \rho_1)$ – количество индивидуумов группы A_1 , доживших от момента времени $t - 1$ до момента времени t , $0 < \rho_1 < 1$ – вероятность дожития от $t - 1$ до t ;

$u_{12}(t) \sim Bin(\hat{x}_1(t), \lambda_{12})$ – количество индивидуумов группы A_1 , перешедших в группу A_2 в течение промежутка времени $(t - 1, t]$, $0 < \lambda_{12} < 1$ – вероятность перехода за $(t - 1, t]$;

$u_{15}(t)$ – количество индивидуумов, переходящих из группы A_1 в группу A_5 в течение промежутка времени $(t - 1, t]$ в результате хотя бы одного контакта с индивидуумами из групп A_5, A_6, A_7, A_8 ;

$u_{15}(t) \sim Bin(\hat{x}_1(t) - u_{12}(t), \mu_{15}(t))$, где

$\mu_{15}(t) = 1 - \prod_{j=5}^8 (1 - r_{1j})^{\hat{x}_j(t)}$ – вероятность инфицирования индивидуума группы A_1 за $(t - 1, t]$;

r_{1j} – вероятность контакта одного индивидуума группы A_1 с одним индивидуумом группы A_j за промежуток времени $(t - 1, t]$, $0 \leq r_{1j} < 1$, $j = 5, \dots, 8$;

$f_1(t) \geq 0$ – случайная величина с заданным законом распределения, отражающая приток индивидуумов в группу A_1 из других регионов за $(t - 1, t]$.

Рассмотрим более подробно вероятность $\mu_{15}(t)$, которую можно интерпретировать как «силу инфекции» (термин из детерминированных моделей). Имеем:

$$\begin{aligned}\mu_{15}(t) &= 1 - \prod_{j=5}^8 (1 - r_{1j})^{\hat{x}_j(t)} = \\ &= 1 - \exp\left(-\sum_{j=5}^8 \hat{x}_j(t) \ln \frac{1}{1 - r_{1j}}\right) \leq \\ &\leq \sum_{j=5}^8 \hat{x}_j(t) \ln \frac{1}{1 - r_{1j}} \approx \sum_{j=5}^8 r_{1j} \hat{x}_j(t).\end{aligned}$$

Выражение вида $\sum_{j=5}^8 r_{1j} \hat{x}_j(t)$ соответствует интенсивности инфицирования индивидуумов группы A_1 в рамках детерминированной модели:

$$\dot{x}_1(t) = -\mu_1 x_1(t) - \gamma_{12} x_1(t) - \sum_{j=5}^8 \alpha_{15j} x_j(t) x_1(t) + \gamma_{21} x_2(t) + f_1(t).$$

Сравните с уравнением в стохастической постановке:

$$x_1(t) = \hat{x}_1(t) - u_{12}(t) - u_{15}(t) + u_{21}(t) + f_1(t).$$

Остальные уравнения строятся в аналогичном виде.

Система уравнений модели:

$$x_1(t) = \hat{x}_1(t) - u_{12}(t) - u_{15}(t) + u_{21}(t) + f_1(t),$$

$$x_2(t) = \hat{x}_2(t) - u_{21}(t) - u_{23}(t) - u_{24}(t) - u_{26}(t) + u_{12}(t) + \\ + u_{32}(t) + u_{42}(t) + f_2(t),$$

$$x_3(t) = \hat{x}_3(t) - u_{32}(t) - u_{37}(t) + u_{23}(t),$$

$$x_4(t) = \hat{x}_4(t) - u_{42}(t) - u_{48}(t) + u_{24}(t),$$

$$x_5(t) = \hat{x}_5(t) + u_{15}(t) + u_{65}(t),$$

$$x_6(t) = \hat{x}_6(t) - u_{65}(t) + u_{26}(t) + u_{76}(t) + u_{86}(t),$$

$$x_7(t) = \hat{x}_7(t) - u_{76}(t) + u_{37}(t),$$

$$x_8(t) = \hat{x}_8(t) - u_{86}(t) + u_{48}(t),$$

$$x_i(0) = x_i^{(0)},$$

$$i = 1, 2, \dots, 8; \quad t = 1, 2, \dots, T.$$

Здесь $(x_1^{(0)}, \dots, x_8^{(0)})$ – целочисленный (неотрицательный) вектор начальных численностей с заданным законом распределения.

Система уравнений на математические ожидания:

$$m_1(t) = (1 - \lambda_{12})\rho_1 m_1(t - 1) + \lambda_{21}\rho_2 m_2(t - 1) - Eu_{15}(t) + Ef_1(t),$$

$$m_2(t) = \lambda_{12}\rho_1 m_1(t - 1) + (1 - \lambda_{21} - \lambda_{23} - \lambda_{24})\rho_2 m_2(t - 1) + \lambda_{32}\rho_3 m_3(t - 1) + \lambda_{42}\rho_4 m_4(t - 1) - Eu_{26}(t) + Ef_2(t),$$

$$m_3(t) = \lambda_{23}\rho_2 m_2(t - 1) + (1 - \lambda_{32})\rho_3 m_3(t - 1) - Eu_{37}(t),$$

$$m_4(t) = \lambda_{24}\rho_2 m_2(t - 1) + (1 - \lambda_{42})\rho_4 m_4(t - 1) - Eu_{48}(t),$$

$$m_5(t) = \rho_5 m_5(t - 1) + \lambda_{65}\rho_6 m_6(t - 1) + Eu_{15}(t),$$

$$m_6(t) = (1 - \lambda_{65})\rho_6 m_6(t - 1) + \lambda_{76}\rho_7 m_7(t - 1) + \lambda_{86}\rho_8 m_8(t - 1) + Eu_{26}(t),$$

$$m_7(t) = (1 - \lambda_{76})\rho_7 m_7(t - 1) + Eu_{37}(t),$$

$$m_8(t) = (1 - \lambda_{86})\rho_8 m_8(t - 1) + Eu_{48}(t),$$

$$m_i(0) = Ex_i^{(0)},$$

$$i = 1, 2, \dots, 8; \quad t = 1, 2, \dots, T.$$

Применяем неравенство Йенсена и метод монотонных операторов.

Подсистема верхних оценок для $m_1(t), m_2(t), m_3(t), m_4(t)$:

$$\begin{pmatrix} m_1(t) \\ m_2(t) \\ m_3(t) \\ m_4(t) \end{pmatrix} \leq \begin{pmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & 0 \\ 0 & b_{42} & 0 & b_{44} \end{pmatrix} \cdot \begin{pmatrix} m_1(t-1) \\ m_2(t-1) \\ m_3(t-1) \\ m_4(t-1) \end{pmatrix} + \begin{pmatrix} Ef_1(t) \\ Ef_2(t) \\ 0 \\ 0 \end{pmatrix},$$

$$b_{11} = (1 - \lambda_{12})\rho_1, \quad b_{12} = \lambda_{21}\rho_2, \quad b_{21} = \lambda_{12}\rho_1,$$

$$b_{22} = (1 - \lambda_{21} - \lambda_{23} - \lambda_{24})\rho_2, \quad b_{23} = \lambda_{32}\rho_3, \quad b_{24} = \lambda_{42}\rho_4,$$

$$b_{32} = \lambda_{23}\rho_2, \quad b_{33} = (1 - \lambda_{32})\rho_3, \quad b_{42} = \lambda_{24}\rho_2, \quad b_{44} = (1 - \lambda_{42})\rho_4.$$

Модули собственных чисел матрицы $B = (b_{ij})$ меньше 1.

Решения $m_1(t), \dots, m_4(t)$ первой подсистемы можно оценить сверху константами $\tilde{m}_1, \dots, \tilde{m}_4$ и подставить эти оценки в уравнения второй подсистемы:

$$\begin{pmatrix} m_5(t) \\ m_6(t) \\ m_7(t) \\ m_8(t) \end{pmatrix} \leq C \cdot \begin{pmatrix} m_5(t-1) \\ m_6(t-1) \\ m_7(t-1) \\ m_8(t-1) \end{pmatrix},$$

$$C = \begin{pmatrix} P_1 Q_{15} + \rho_5 & P_1 Q_{16} + \lambda_{65} \rho_6 & P_1 Q_{17} & P_1 Q_{18} \\ P_2 Q_{25} & P_2 Q_{26} + (1 - \lambda_{65}) \rho_6 & P_2 Q_{27} + \lambda_{76} \rho_7 & P_2 Q_{28} + \lambda_{86} \rho_8 \\ P_3 Q_{35} & P_3 Q_{36} & P_3 Q_{37} + (1 - \lambda_{76}) \rho_7 & P_3 Q_{38} \\ P_4 Q_{45} & P_4 Q_{46} & P_4 Q_{47} & P_4 Q_{48} + \rho_8 \end{pmatrix},$$

$$P_1 = (1 - \lambda_{12}) \rho_1 \tilde{m}_1, \quad P_2 = (1 - \lambda_{21} - \lambda_{23} - \lambda_{24}) \rho_2 \tilde{m}_2,$$

$$P_3 = (1 - \lambda_{32}) \rho_3 \tilde{m}_3, \quad P_4 = (1 - \lambda_{42}) \rho_4 \tilde{m}_4,$$

$$Q_{ij} = \rho_j \ln \frac{1}{1 - r_{ij}}, \quad i = 1, \dots, 4, \quad j = 5, \dots, 8.$$

Достаточное условие асимптотической устойчивости нулевого решения

(= критерий асимптотической устойчивости системы верхних оценок)

Элементы матрицы C неотрицательны.

Если главные миноры матрицы $I - C$ положительны (т.е. $I - C$ является невырожденной М-матрицей), то $m_i(t) \rightarrow 0$ при $t \rightarrow \infty$, $i = 5, 6, 7, 8$.

Выполнение этого критерия для матрицы $I - C$ обеспечивает полное искоренение инфекции: с течением времени численности групп A_5, A_6, A_7, A_8 стремятся к нулю с вероятностью 1.

Вычислительный эксперимент

Приведены статистические оценки математических ожиданий общей численности восприимчивых ($x_1(t) + x_2(t) + x_3(t) + x_4(t)$) и инфицированных ($x_5(t) + x_6(t) + x_7(t) + x_8(t)$) индивидуумов для двух наборов параметров.

Для первого набора параметров $I - C$ является невырожденной M -матрицей, что обеспечивает искоренение инфекции (эксперимент 1).

Для второго набора параметров указанный для $I - C$ критерий не выполняется, оценка математического ожидания количества инфицированных индивидуумов с течением времени выходит на ненулевой стационарный уровень (эксперимент 2).

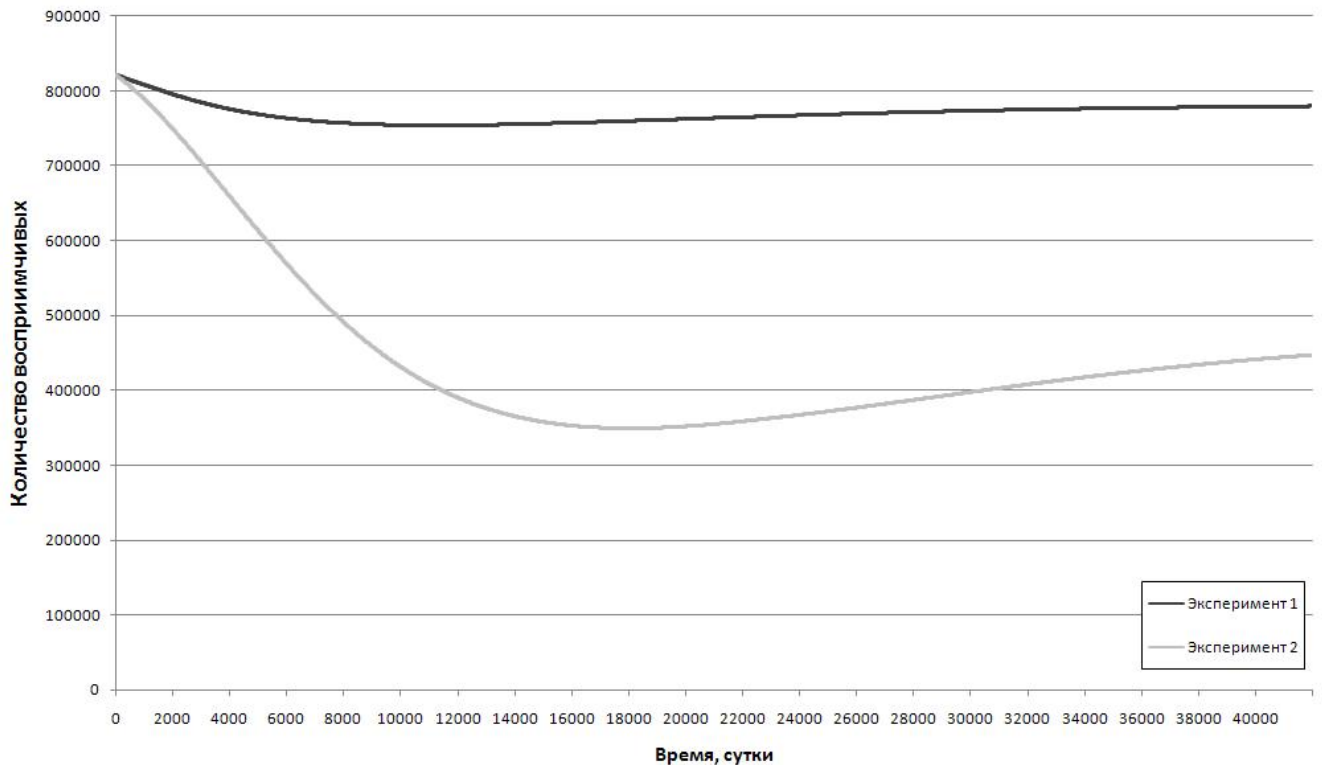


Рис. 1. Динамика среднестатистической численности индивидуумов, восприимчивых к ВИЧ-инфекции.

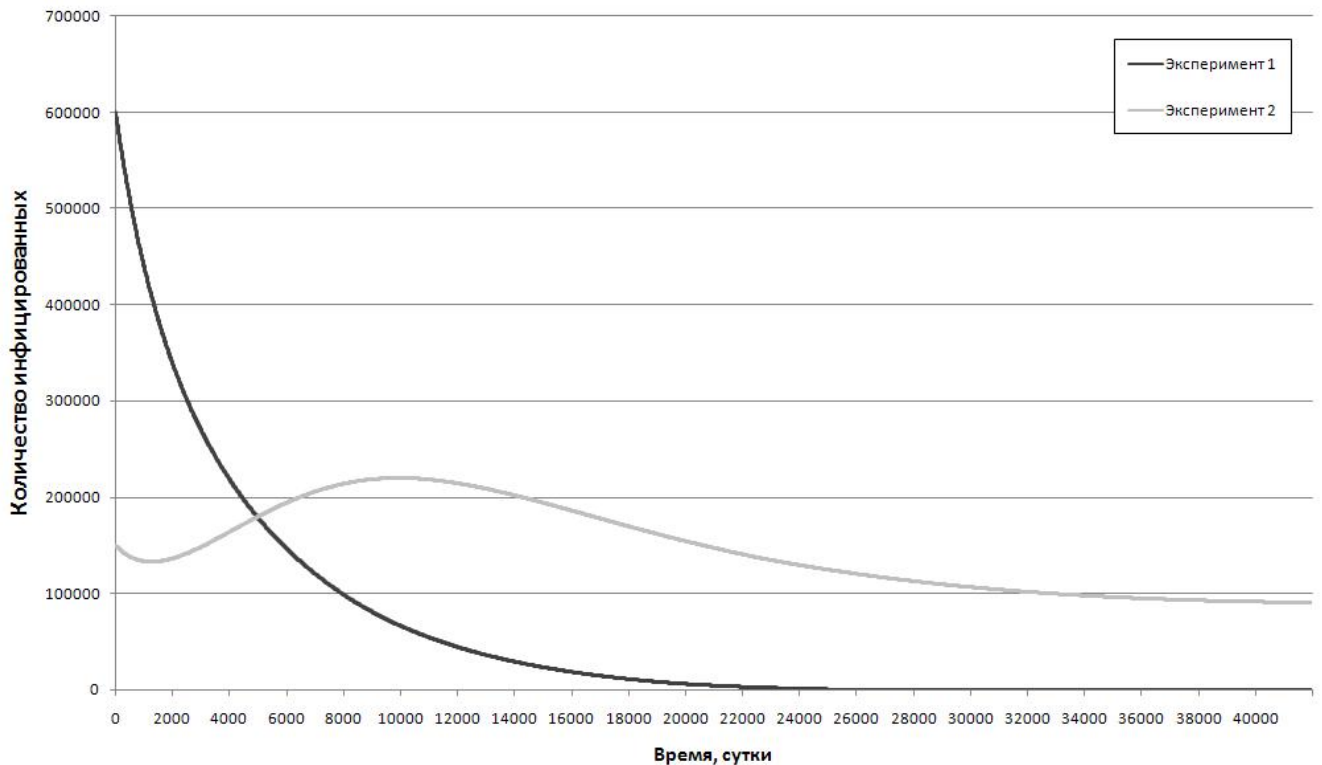


Рис. 2. Динамика среднестатистической численности ВИЧ-инфицированных индивидуумов.

Некоторые обобщения

При моделировании популяционной динамики обычно используются следующие зависимости:

size dependent, stage dependent, age dependent.

Для учета указанных факторов в моделях распространения эпидпроцессов можно брать переменные вида

$x_{i,j,k}(t)$ — численность индивидуумов группы i , возраста j , находящихся в стадии k .

Проблема быстродействия моделирующих программ

- Разработка параллельных алгоритмов для высокопроизводительных вычислительных систем (MPI, OpenMP)
- Распределённые вычисления на GPU NVIDIA
- Переход к тензорным вычислениям

Спасибо за внимание!