На правах рукописи

ВИНОГРАДОВА Марина Станиславовна

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ РАЗВИТИЯ ИЗОЛИРОВАННОЙ КЛЕТОЧНОЙ ПОПУЛЯЦИОННОЙ СИСТЕМЫ

Специальность 05.13.01 – Системный анализ, управление и обработка информации (информатика, машиностроение)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2013

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Московский государственный технический университет имени Н.Э. Баумана»

Научный руководитель: доктор физико-математических наук, профессор **Волков Игорь Куприянович**

Официальные оппоненты: **Формалев Владимир Федорович,** заслуженный деятель науки РФ, доктор физико-математических наук, профессор, федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Московский авиационный институт (национальный исследовательский университет)», профессор кафедры вычислительной математики и программирования

Тимонин Владимир Иванович, доктор физико-математических наук, доцент, федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Московский государственный технический университет имени Н.Э. Баумана», профессор кафедры высшей математики

Ведущая организация: федеральное государственное бюджетное учреждение науки «Институт вычислите математики Российской академии наук»

Защита диссертации состоится «15» октября 2013 г. в 13 ч. на заседании диссертационного совета Д 212.141.15 при Московск дарственном техническом университете имени Н.Э. Баумана по адресу: Рубцовская наб., 2/18, ауд. 1006 л.

С диссертацией можно ознакомиться в библиотеке Мос государственного технического университета имени Н.Э. Баумана.

Автореферат разослан «12» сентября 2013 г.

Ученый секретарь диссертационного совета, кандидат технических наук, доцент А.В. Аттетков

НТБ МГТУ им. Н.Э. Баумана
2034313
Виноградова М.С. Математи

Общая характеристика работы

Актуальность темы. Современные информационные технологии, основанные на использовании методов математического моделирования и методов обработки информации, полученной на основе натурных экспериментов, во многих случаях позволяют получить такое описание процессов функционирования сложной системы, на основании которого возможен анализ дальнейших сценариев ее развития и принятие решения о вмешательстве с целью исключения нежелательных последствий. Так, активно разрабатываются информационные системы медицинского назначения, предназначенные для формализации процесса постановки диагноза при различных заболеваниях и контроля процесса лечения.

В последнее десятилетие проводятся широкие исследования стволовых клеток с целью изучения возможностей применения этих клеток в регенеративной медицине и клеточной терапии (Н.П. Бочков; 2008).

Предназначенные для пересадки клетки получают в нужном количестве путем их культивирования в лабораторных условиях, поскольку исходного количества клеток, взятых из организма пациента, недостаточно.

В процессе размножения в силу естественной изменчивости в культуре клеток могут появиться клетки с хромосомными мутациями (аномальные клетки), причем клоны аномальных клеток могут обладать селективным преимуществом по сравнению с нормальными клетками (Н.П. Бочков; 2008). При трансплантации такого материала имеется риск возникновения канцерогенеза у пациента, поэтому остро встает проблема безопасности при проведении клеточной терапии у широкой категории пациентов (Е.Ю. Осипова; 2009).

При этом возникает необходимость разработки критериев «отбраковывания» культур, в которых идет преимущественное размножение клеток с аномальными хромосомными наборами (клонообразование) с подавлением развития нормальных клеток.

Выявление факта преобладания аномальных клеток на поздних стадиях выращивания ведет к необходимости повторения процесса.

Для отбраковки культуры клеток на ранних стадиях развития и анализа процессов клонообразования возможно использование современных информационных технологий, основанных на методах математического моделирования и параметрической идентификации моделей на ограниченных выборках экспериментальных данных о значениях вектора состояния изучаемой системы, а также статистических методах. Для решения указанной задачи необходимо разработать математические модели развития клеточных популяций и адаптировать к особенностям разрабатываемых математических моделей известные алгоритмы их параметрической идентификации на ограниченных выборках, а также найти законы распределения оценок параметров моделей.

М Г Т У
им Н Э Б ауна
Б иблиотека

Использование методов математического моделирования в биологии и медицине получает все более широкое распространение. Здесь можно выделить различные модели межвидовой конкуренции в биологии(Г.Ю. Ризниченко, А.В. Рубин; 1993), модели различных заболеваний человека (А.А. Романюха; 2011) или процессов, происходящих в организме, в том числе при онкологических заболеваниях (Massey S.C.; 2012), а также модели развития клеточных популяций и популяций микроорганизмов, ориентированные на оценку влияния среды на рост популяции (Ю.М. Романовский; 1984). Среди известных моделей можно отметить модели роста опухолей, учитывающие иммунный ответ организма и воздействие химиотерапии (M.R. Owen; 1997).

Необходимость разработки различных математических моделей, описывающих как процессы развития клеточных популяций, так и процессы, протекающие в самой клетке, подчеркивается многими авторами.

Известные современные математические модели развития различных клеточных популяций (Y. Luo, C.L. Lim; 2012), базируются на методах стохастического моделирования, что не позволяет провести их параметрическую идентификацию.

В динамической модели развития популяций здоровых и раковых клеток (Y.C. Ou; 2001) учитываются вероятность перерождения нормальных клеток в раковые и селективное преимущество, которым обладают раковые клетки. При этом не рассматривается межпопуляционное взаимодействие.

Методы параметрической идентификации математической модели, позволяющие получить оценки параметров модели с использованием достаточно малых массивов измерений, в настоящее время известны (И.К. Волков; 1994), однако для конкретных моделей необходима их доработка и адаптация с учетом особенностей моделей.

Один из подходов к получению законов распределения оценок параметров математических моделей базируется на байесовском подходе и теории инвариантности (H. Jeffreys; 1983). Указанный подход активно применяется в задачах параметрической идентификации математических моделей экономических процессов (А. Зельнер; 1980).

Знание законов распределения, а также точечных и интервальных оценок для параметров модели позволяет методами математического моделирования получить оценки вероятностей реализации различных сценариев дальнейшего развития изучаемой клеточной популяционной системы.

Основная цель исследований состоит в проведении системного анализа процесса развития изолированной клеточной популяционной системы в лабораторных условиях, разработке математических моделей функционирования этой системы и решении задачи их параметрической идентификации на ограниченных выборках экспериментальных данных.

Для достижения поставленной цели потребовалось решение **следующих основных задач**:

— проведение системного анализа процесса развития изолированной клеточной популяции в лабораторных условиях и выявление его основных «механизмов»;

— построение математических моделей динамики развития изолированной клеточной популяционной системы в лабораторных условиях;

— параметрический анализ разработанных моделей с привлечением аналитических и численных методов;

— разработка алгоритма идентификации параметров моделей по ограниченным массивам данных наблюдений;

— теоретический анализ законов распределения оценок параметров модели;

— получение точечных и интервальных оценок для параметров математических моделей развития популяции.

Методы исследования. При решении задач, возникших в ходе выполнения диссертационной работы, использовались различные классы математических методов: системного анализа, теории дифференциальных уравнений, теории устойчивости, теории вероятностей и математической статистики.

Научная новизна. В диссертации получены следующие научные результаты, которые выносятся на защиту.

1. Линейная математическая модель развития взаимодействующих клеточных популяций нормальных и аномальных стволовых клеток при стандартных лабораторных условиях культивирования при отсутствии конкуренции за ресурсы.

2. Нелинейная математическая модель развития взаимодействующих клеточных популяций нормальных и аномальных стволовых клеток в лабораторных условиях, учитывающая ограничения на ресурсы.

3. Условия существования положений равновесия нелинейной модели, зависимости их характера от параметров модели.

4. Теоретические функции плотности распределения вероятностей параметров линейной и нелинейной математических моделей, полученные на основе байесовского подхода и теории инвариантности Джеффриса.

5. Маргинальные функции плотности распределения вероятностей параметров математической модели, интервальные оценки этих параметров.

Достоверность и обоснованность научных результатов подтверждена строгими математическими доказательствами и результатами численного моделирования. Результаты диссертационной работы согласуются с известными результатами других авторов.

Практическая ценность полученных результатов состоит в том, что разработанные математические модели позволяют провести качественный анализ возможных сценариев развития клеточных популяций и могут быть использованы

для анализа последствий различных внешних воздействий на развитие клеточных популяций.

Теоретические результаты доведены до конструктивных методов, позволяющих для разработанных математических моделей получать точечные и интервальные оценки параметров разработанных моделей. Указанные оценки могут быть использованы для анализа вероятностей реализации различных сценариев развития клеточных популяций.

Апробация результатов работы. Результаты работы докладывались на XVIII-й Международной конференции «Математика. Компьютер. Образование» (Пущино, 2011), VIII Международной конференции «Молекулярная генетика соматических клеток» (Звенигород, 2011), научном семинаре кафедры «Математическое моделирование» МГТУ им. Н.Э. Баумана (2013).

Диссертация является составной частью фундаментальных исследований, проводимых в рамках грантов № 09-04-00948, № 10-07-00468, № 12-07-00329, № 13-07-00720 и проекта 2.1.1/227 аналитической ведомственной программы «Развитие научного потенциала высшей школы на 2009-2011 гг.».

Публикации. Основные научные результаты отражены в 6 статьях в научных журналах, которые включены в Перечень ведущих научных журналов и изданий для опубликования основных научных результатов диссертации, рекомендованных ВАК РФ, и 2 тезисах докладов. Общий объем 5,56 п.л.

Личный вклад соискателя. Все исследования, результаты которых изложены в диссертации, проведены лично соискателем в процессе научной деятельности. Из совместных работ в диссертацию включен лишь тот материал, который непосредственно принадлежит соискателю, заимствованный материал обозначен в работе ссылками.

Структура и объем работы. Диссертация состоит из введения, четырех глав, выводов и списка литературы. Работа изложена на 128 страницах, содержит 25 иллюстраций. Библиография включает 95 наименований.

Содержание работы

Во **введении** обоснована актуальность темы, сформулированы цели и задачи исследования, научная новизна, теоретическая и практическая значимость полученных результатов, основные положения, выносимые на защиту, а также приведены данные о структуре и объеме диссертации.

В **первой главе** проведен системный анализ факторов, определяющих развитие популяционной системы, состоящей из двух видов стволовых клеток человека: нормальных и аномальных, при их культивировании в лабораторных условиях.

На основе проведенного анализа предложена линейная модель динамики взаимодействующих клеточных популяций.

В разделе 1.1 приводятся исходные допущения, принятые при построении модели. Эти допущения справедливы при культивировании стволовых клеток в условиях отсутствия конкуренции за ресурсы.

Аномальными будем называть клетки, имеющие численные аномалии хромосомного набора. Нормальными считаем все остальные клетки в популяционной системе.

При рассмотрении достаточно большой совокупности стволовых клеток, образующих изолированную популяцию, можно говорить о средней продолжительности клеточного цикла как временной характеристике процессов, происходящих в данной популяции.

При этом для популяции нормальных клеток можно говорить о доле клеток, которая за время среднего клеточного цикла разделилась на пару нормальных клеток, доле, которая разделилась на пару аномальных клеток, доле клеток, которые не разделились и остались живыми, а также о доле клеток, которые погибли. Аналогично аномальные клетки за время среднего клеточного цикла могут выжить или погибнуть, выжившие клетки могут не разделиться, или разделиться, оставаясь при этом аномальными.

При построении математической модели сделаны следующие допущения:

1) рассматриваемая клеточная популяционная система является изолированной и имеет значительную численность, что позволяет использовать детерминированную модель ее развития;

2) ресурсы считаются неограниченными, поскольку культивирование проводится в стандартных лабораторных условиях в культуральных флаконах при регулярном «пересеве» клеток со сменой части (или всего объема) питательной среды;

3) внешние воздействия на популяционную систему имеют квазистационарный характер, что позволяет считать постоянными параметры модели;

4) изменение количества нормальных и аномальных клеток возможно либо в процессе клеточного деления либо в процессе естественной гибели клеток.

При построении модели в качестве исходных были приняты следующие параметры: τ^0, τ^1 — средние продолжительности клеточных циклов в популяции нормальных и популяции аномальных клеток; A^0, A^1 — доли клеток в популяции нормальных и популяции аномальных клеток, «погибающих» на временных интервалах длительности τ^0 и τ^1 соответственно; M^0, M^1 — доли клеток в указанных популяциях, разделившихся на временных интервалах длительности τ^0 и τ^1 соответственно; γ^0 — доля нормальных клеток, «переходящих» на интервале длительности τ^0 в процессе деления в популяцию аномальных клеток. Считаем, что $0 < A^0 < 1, 0 < A^1 < 1, 0 < M^0 < 1, 0 < M^1 < 1$ и $0 < \gamma^0 < 1$.

В разделе 1.2 получена математическая модель развития популяционной системы

$$\frac{dx}{dt} = \tilde{q}_{00}x, \quad \frac{dy}{dt} = \tilde{q}_{10}x + \tilde{q}_{11}y. \tag{1}$$

Здесь $x = X/X_0$, $y = Y/X_0$, где X и Y — текущие численности популяций нормальных и аномальных клеток, X_0 — численность популяции нормальных клеток в начальный момент времени, $t = \tau/\tau^0$ — безразмерное модельное время, τ — реальное время, $\tilde{q}_{00} = (1 - A^0)(1 + M^0 - 2M^0\gamma^0) - 1$, $\tilde{q}_{10} = 2(1 - A^0)M^0\gamma^0$, $\tilde{q}_{11} = \mu(-A^1 + M^1 - A^1M^1)$, а $\mu = \tau^0/\tau^1$.

В разделе 1.3 проведен анализ сценариев развития клеточной популяционной системы, реализующихся при различных значениях параметров системы (1). Показано, что все решения системы (1), начинающиеся в множестве $G = \{(x, y) : x \geq 0,\ y \geq 0\}$, не выходят из этого множества.

Основные сценарии развития клеточной популяционной системы определяются значениями параметров $\tilde{q}_{00}$ и $\tilde{q}_{11}$. Благоприятные сценарии развития реализуются при $\tilde{q}_{00} > 0$, $\tilde{q}_{11} > 0$, $\tilde{q}_{00} > \tilde{q}_{11}$, а также при $\tilde{q}_{00} > 0$, $\tilde{q}_{11} < 0$. При значениях параметров из указанных множеств соотношение численностей нормальных и аномальных клеток с течением времени стремится к постоянной величине.

Неблагоприятные сценарии развития, при которых популяция нормальных клеток подавляется популяцией аномальных клеток, реализуются при $\tilde{q}_{00} > 0$, $\tilde{q}_{11} > 0$, $\tilde{q}_{00} < \tilde{q}_{11}$ и при $\tilde{q}_{00} < 0$, $\tilde{q}_{11} > 0$.

Также неблагоприятным является сценарий полного вымирания популяционной системы, реализующийся при $\tilde{q}_{00} < 0$, $\tilde{q}_{11} < 0$.

Во **второй главе** на основе математической модели, предложенной в главе 1, построена нелинейная модель динамики развития популяционной системы, учитывающая ограниченность ресурсов.

В разделе 2.1 предложена математическая модель, в которой учтена функциональная зависимость параметров M^0 и M^1 от суммарной численности популяций нормальных ($j = 0$) и аномальных ($j = 1$) клеток. Пусть

$$f_0(x, y) = \alpha_0 - \beta_{00}x - \beta_{01}y, \quad f_1(x, y) = \alpha_1 - \beta_{10}x - \beta_{11}y. \tag{2}$$

Положим

$$M^j(x, y) = \begin{cases} 0 & \text{при } f_j(x, y) \leq 0; \\ 1 & \text{при } f_j(x, y) \geq 1; \\ f_j(x, y) & \text{иначе,} \end{cases} \tag{3}$$

где $j \in \{0, 1\}$. Параметры $\beta_{ij} > 0$, $i, j \in \{0, 1\}$, задают уменьшение доли делящихся клеток как за счет внутрипопуляционной ($j = i$), так и за счет межпопуляционной ($j \neq i$) конкуренции. Интерпретация параметра α_j идентична интерпретации параметра M^j, и $0 < \alpha_j < 1$, $j \in \{0, 1\}$.

Модель развития популяционной системы имеет вид

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= \left((1-A^0)(1+M^0(x,y)-2M^0(x,y)\gamma^0)-1\right)x,\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= \mu\left(-A^1+M^1(x,y)-A^1M^1(x,y)\right)y+\\ &\qquad +2M^0(x,y)(1-A^0)\gamma^0 x,\end{aligned} \tag{4}$$

Показано, что траектории системы (4), начинающиеся в точках множества $G=\{(x,y): x\geq 0,\ y\geq 0\}$ не выходят из этого множества.

В множестве G в общем случае можно выделить четыре подмножества

$$\begin{aligned}G_1&=\{(x,y): x\geq 0,\ y\geq 0,\ f_0(x,y)\geq 0,\ f_1(x,y)\geq 0\},\\ G_2&=\{(x,y): x\geq 0,\ y\geq 0,\ f_0(x,y)\leq 0,\ f_1(x,y)\leq 0\},\\ G_3&=\{(x,y): x\geq 0,\ y\geq 0,\ f_0(x,y)\leq 0,\ f_1(x,y)\geq 0\},\\ G_4&=\{(x,y): x\geq 0,\ y\geq 0,\ f_0(x,y)\geq 0,\ f_1(x,y)\leq 0\},\end{aligned} \tag{5}$$

на каждом из которых система (4) имеет различный вид.

Так, на множестве G_1

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= h_{01}x+h_{02}x^2+h_{03}xy,\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= h_{11}x+h_{12}y+h_{13}x^2+h_{14}xy+h_{15}y^2,\end{aligned} \tag{6}$$

где

$$\begin{aligned}h_{01}&=(1-A^0)(1-2\gamma^0)\alpha_0-A^0, \quad h_{02}=-\beta_{00}(1-A^0)(1-2\gamma^0),\\ h_{03}&=-\beta_{01}(1-A^0)(1-2\gamma^0), \quad h_{11}=2\alpha_0(1-A^0)\gamma^0,\\ h_{12}&=\mu((1-A^1)\alpha_1-A^1), \quad h_{13}=-2\beta_{00}(1-A^0)\gamma^0,\\ h_{14}&=-(\mu\beta_{10}(1-A^1)+2\beta_{01}(1-A^0)\gamma^0), \quad h_{15}=-\mu\beta_{11}(1-A^1).\end{aligned} \tag{7}$$

Также в разделе приведены виды системы (4) на множествах G_2, G_3 и G_4 и описаны варианты взаимного расположения множеств G_1, G_2, G_3 и G_4.

В разделе 2.2 выполнено исследование поведения траекторий. Доказано следующее утверждение: если параметры нелинейной системы (4) удовлетворяют ограничениям $0<A^j<1$, $0<\gamma^0<1$, $0<\alpha_j<1$, $0<\beta_{ij}<1$, $i,j\in\{0,\ 1\}$, то все траектории системы (4), начинающиеся в множестве $G_{134}=G_1\cup G_3\cup G_4$, не выходят из этого множества.

В разделе 2.3 исследованы точки покоя системы (6) на множестве G_1, которые разделены на вырожденные (хотя бы одна из координат точки является нулевой) и внутренние (обе координаты — положительные числа).

Теоретически обосновано, что в общем случае изучаемая популяционная система не может иметь более двух вырожденных точек покоя в области G_1. При этом точка покоя $O_1(0,0)$ соответствует полному вымиранию популяционной системы, а точка покоя $O_2(0,-h_{12}/h_{15})$ соответствует ее вырождению в популяцию аномальных клеток предельно возможной численности.

При $\gamma^0 = 1/2$ популяционная система имеет лишь вырожденные точки покоя, и для определения ее внутренних точек покоя полагаем $\gamma^0 \in (0.1/2) \cup (1/2,1)$. Показано, что система не может иметь более двух внутренних точек покоя, координаты x которых определяются как положительные решения квадратного уравнения

$$Ax^2 + Bx + C = 0, \tag{8}$$

где $A = h_{13} - h_{14}k + h_{15}k^2$, $B = h_{11} - h_{12}k + h_{14}b - 2h_{15}bk$, $C = h_{12}b + h_{15}b^2$. Вторая координата находится из соотношения $y = -kx + b$, где

$$k = \frac{h_{02}}{h_{03}} = \frac{\beta_{00}}{\beta_{01}}, \quad b = -\frac{h_{01}}{h_{03}} = \frac{1}{\beta_{01}}\left\{\alpha_0 - \frac{A^0}{(1-A^0)(1-2\gamma^0)}\right\},$$

и из условия $y \geq 0$. Необходимым условием существования внутренних точек покоя является выполнение условия $b > 0$. Если внутренняя точка покоя (x,y), где $x > 0$, $y > 0$, найдена, необходимо проверить принадлежность этой точки множеству G_1.

В разделе 2.4 исследованы условия существования точек покоя в множествах G_2, G_3 и G_4. Установлено, что в множестве G_2 точек покоя не существует, в множестве G_3 при некоторых значениях параметров может существовать вырожденная точка покоя, а в множестве G_4 не может быть вырожденных точек покоя, но может существовать не более двух внутренних точек.

В разделе 2.5 проведено исследование устойчивости вырожденных точек покоя по первому приближению и проанализированы основные критические случаи.

В разделе 2.6 проведен анализ результатов, полученных в разделе 2.5, и получены фазовые портреты, соответствующие различным сценариям развития популяции. Основные результаты удобно записать с использованием параметров

$$d_0 = \alpha_0 - \frac{A^0}{(1-A^0)(1-2\gamma^0)}, \quad d_1 = \alpha_1 - \frac{A^1}{1-A^1}.$$

Если $\gamma^0 \in (0,1/2)$, то точка покоя O_1 будет устойчивым узлом при $d_0 < 0$ и $d_1 < 0$, неустойчивым узлом при $d_0 > 0$ и $d_1 > 0$, седлом при выполнении условий $d_0 > 0$ и $d_1 < 0$ или $d_0 < 0$ и $d_1 > 0$.

Точка покоя O_2 будет устойчивым узлом при выполнении условий $d_0 < d_1\beta_{01}/\beta_{11}$ и $d_1 > 0$, седлом при $d_0 > d_1\beta_{01}/\beta_{11}$ и $d_1 > 0$. Точка покоя O_2 не может быть неустойчивым узлом.

При $\gamma^0 \in (1/2, 1)$ точка покоя O_1 будет устойчивым узлом при $d_0 > 0$ и $d_1 < 0$, седлом при выполнении условий $d_0 > 0$ и $d_1 > 0$.

Точка покоя O_2 будет устойчивым узлом при $d_0 > d_1\beta_{01}/\beta_{11}$ и $d_1 > 0$, седлом при $d_0 < d_1\beta_{01}/\beta_{11}$ и $d_1 > 0$.

При указанном диапазоне γ^0 точки покоя O_1 и O_2 не могут быть неустойчивыми узлами.

Исследование устойчивости внутренних точек покоя ввиду громоздкости получающихся выражений возможно, по-видимому, только численно.

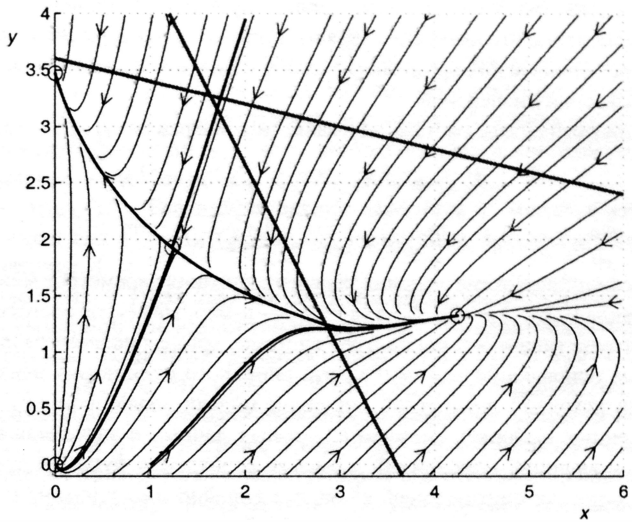


Рис. 1. Пример фазового портрета

Фазовый портрет системы (4), приведенный на Рис. 1, получен при следующих значениях параметров модели: $A^0 = 0.200$, $A^1 = 0.275$, $\gamma^0 = 0.150$, $\alpha_0 = 0.900$, $\alpha_1 = 0.900$, $\beta_{00} = 0.050$, $\beta_{01} = 0.250$, $\beta_{10} = 0.250$, $\beta_{11} = 0.150$, $\mu = 1$.

При указанных значениях параметров имеются четыре точки покоя. Вырожденная точка с координатами $(0, 0)$ является неустойчивым узлом, а вторая вырожденная точка покоя — устойчивым узлом. Имеются две внутренние точки покоя — седло, расположенное в множестве G_1, и устойчивый узел, расположенный в множестве G_4.

На рисунке проведены границы множеств G_i и сепаратрисы, которые ограничивают область, из точек которой траектории стремятся к устойчивому узлу. Приведена траектория, выходящая из точки $(1, 0)$. Данный вариант является благоприятным для развития популяции нормальных клеток при начале траектории в окрестности этой точки.

В **третьей главе** решена задача параметрической идентификации линейной модели, предложенной в главе 1, с использованием ограниченных выборок экспериментальных данных.

В разделе 3.1 рассмотрена задача оценивания параметров модели. Непосредственно оценить значения параметров указанной модели не удается, поскольку по результатам экспериментов можно оценить только количество нормальных и аномальных клеток по различным пассажам (Н.П. Бочков; 2008)

Соответствующие измерения количества нормальных и аномальных клеток можно считать проведенными через равные промежутки времени Δt.

Будем считать, что измерения являются прямыми и содержат случайные ошибки, распределенные по нормальному закону с нулевом математическим ожиданием и неизвесной дисперсией.

Дискретный аналог системы (1), записанный в реальном времени, примет вид

$$X_{k+1} = a^0 X_k, \quad Y_{k+1} = a^1 Y_k + b^1 X_k, \tag{9}$$

где $a^0 = \frac{\tilde{q}_{00}}{\tau^0}\Delta t + 1$, $a^1 = \frac{\tilde{q}_{11}}{\tau^1}\Delta t + 1$, $b^1 = \frac{\tilde{q}_{10}}{\tau^1}\Delta t$.

Описана общая схема решения задачи оценивания параметров математической модели вида $V(t_{k+1}) = AV(t_k)$, где $V(t_k)$ — n-мерный вектор состояния системы, A — матрица типа $n \times n$ параметров системы, которая полагается неизвестной.

Схема основана на том, что задача линейного оценивания неизвестных параметров исходной дискретной линейной модели эволюционного процесса на ограниченных выборках экспериментальных данных о значениях векторов состояния и внешних воздействий в смысле совпадения функций правдоподобия эквивалентна соответствующей ей задаче линейного оценивания с ошибками в инструментальных переменных, записанной на максимально возможной системе непересекающихся шаблонов разностного уравнения математической модели изучаемого эволюционного процесса (И.К. Волков; 1994).

Указаны условия, при которых рассматриваемая задача линейного оценивания с ошибками в инструментальных переменных может быть представлена как задача регрессионного анализа. Приведены условия, при выполнении которых оценка максимального правдоподобия для матрицы неизвестных параметров модели в рассматриваемом случае совпадает с ее оценкой, получаемой методом наименьших квадратов. Описан алгоритм решения задачи параметрической идентификации математической модели, включающий нахождение точечных оценок параметров модели, проверку статистической гипотезы о виде закона распределения случайных возмущений, нахождение законов распределения параметров модели и построения доверительных интервалов для параметров модели.

В разделе 3.2 приведены основные принципы байесовского подхода к задаче получения точечных оценок параметров модели. Этот подход, основанный на

применении теоремы Байеса, часто используется при небольших объемах данных и неизвестной ковариционной матрице Σ. При его реализации неизвестные параметры модели рассматриваются как случайные величины. Также в разделе приведены основные положения теории инвариантности Джеффриса, на основе которой были выбраны априорные функции плотности распределения вероятностей параметров модели.

В разделе 3.3 обоснован метод получения точечных оценок параметров линейной модели. Показано, что в силу специфики модели задачу идентификации матрицы параметров системы можно разбить на две независимые подзадачи: нахождение оценки параметра a^0 и нахождение оценок параметров a^1 и b^1.

Пусть $\{t_{2n-1}; t_{2n}\}$ — n-ый двухточечный шаблон первого уравнения системы (9), где $t_{2n-1} = (2n-1)\Delta t$, $t_{2n} = 2n\Delta t$, в узлах которого известны экспериментальные значения X_{2n-1}, X_{2n} суммарной численностей популяций нормальных клеток в моменты фиксации t_{2n-1} и t_{2n} соответственно, а массивы экспериментальных данных имеют вид: $X_{ч} \triangleq (X_2, \ldots, X_{2N})$, $X_{нч} \triangleq (X_1, \ldots, X_{2N-1})$, $X_{ч}, X_{нч} \in M_{1\times N}(R)$. Реализация оценки параметра a^0 методом наименьших квадратов имеет вид $\hat{a}^0 = X_{ч}X^{+}_{нч}$, где $X^{+}_{нч}$ — матрица, псевдообратная по отношению к матрице $X_{нч}$.

Для известных в узлах того же шаблона экспериментальных значений Y_{2n-1}, Y_{2n} суммарной численностей популяций аномальных клеток массивы экспериментальных данных задаются в виде: $Y_{ч} \triangleq (Y_2, \ldots, Y_{2N}) \in M_{1\times N}(R)$, $Y_{нч} \triangleq (Y_1, \ldots, Y_{2N-1}) \in M_{1\times N}(R)$, $Z \triangleq (Y^{\mathrm{T}}_{нч}, X^{\mathrm{T}}_{нч})^{\mathrm{T}} \in M_{2\times N}(R)$, а реализация оценки параметров a^1, b^1 методом наименьших квадратов находятся по формуле $(\hat{a}^1, \hat{b}^1) = Y_{ч}Z^{+}$, где Z^{+} — матрица, псевдообратная по отношению к матрице Z.

В разделе 3.4 приведена методика проверки статистической гипотезы о виде закона распределения случайных возмущений. Для решения этой задачи использован критерий А.Н. Колмогорова в модификации Ю.Н. Тюрина, адаптированный к выборкам малого объема. Описан алгоритм проверки статистической гипотезы о нормальном распределении невязок. Алгоритм проиллюстрирован примером.

В разделе 3.5 решены задачи построения законов распределения параметров и задача интервального оценивания параметров линейной математической модели.

В подразделе 3.5.1 с использованием байесовского подхода получена функция плотности распределения вероятностей параметра a^0 на основе массивов $X_{ч}$, $X_{нч}$ экспериментальных данных. При указанном подходе совместная апостериорная функция плотности распределения вероятностей параметра a^0 и дисперсии σ^2 согласно теореме Байеса может быть представлена в виде

$$f(a^0, \sigma^2 | X_{ч}, X_{нч}) \sim f(a^0, \sigma^2) f(X_{ч} | X_{нч}, a^0, \sigma^2), \tag{10}$$

где $f(X_{ч}|X_{нч}, a^0, \sigma^2)$ — условная плотность распределения вероятностей новых наблюдений $X_{ч}$ при определенных значениях параметра a^0, σ^2 и $X_{нч}$, рассматриваемая как функция правдоподобия, а $f(a^0, \sigma^2)$ — совместная априорная функция

плотности распределения вероятностей величин a^0 и σ^2. Здесь знак $\sim$ обозначает пропорциональность.

Апостериорная функция плотности распределения вероятностей $f(X_ч|X_{нч}, a^0, \sigma^2)$ при нормальном распределении ошибок измерений ε_i, $i = \overline{1, N}$, заданных значениях a^0, $X_{нч}$ и выборочной дисперсии σ^2 имеет вид

$$f(X_ч|X_{нч}, a^0, \sigma^2) \sim (\sigma^2)^{-\frac{N}{2}} \exp[-\tfrac{1}{2\sigma^2}(X_ч - a^0 X_{нч})(X_ч - a^0 X_{нч})^т] \qquad (11)$$

Записав совместную функцию правдоподобия для параметра a^0 и дисперсии σ^2 как плотность вероятности совместного появления результатов выборки $x_k^{нч}$, $x_k^ч$, $k = \overline{1. N}$, получим

$$l(a^0, \sigma^2|X_{нч}, X_ч) \sim (\sigma^2)^{-\frac{N}{2}} \exp\left[-(S^2 - (a^0 - \hat{a}^0)^2(X_{нч}X_{нч}^т))/2\sigma^2\right] \qquad (12)$$

Совместную априорную функцию плотности распределения вероятностей $f(a^0, \sigma^2)$ будем считать расплывчатой (H. Jeffreys; 1983). Также примем допущение, что параметр a^0 и выборочная дисперсия σ^2 независимо распределены.

Априорную функцию плотности распределения вероятностей параметра a^0 $f(a^0)$ будем считать константой, поскольку значении параметра a^0 ничего сказать нельзя, и формально a^0 может принимать значения от $-\infty$ до ∞.

Поскольку $\sigma^2 \in [0, +\infty)$, то априорная функция плотности распределения вероятностей для σ принята в виде $f(\sigma^2) \sim \frac{1}{\sigma}$.

Проведя необходимые преобразования, получим функцию плотности распределения вероятностей параметра a^0

$$f(a^0|X_ч, X_{нч}) \sim \{S^2 + (a^0 - \hat{a}^0)^2(X_{нч}X_{нч}^т)\}^{-\frac{N}{2}}, \qquad (13)$$

где $S^2 \triangleq (X_ч - \hat{a}^0 X_{нч})(X_ч - \hat{a}^0 X_{нч})^т \equiv \hat{\varepsilon}\hat{\varepsilon}^т$.

В подразделе 5.3.2 на основе байесовского подхода построена совместная функция плотности распределения вероятностей параметров a^1 и b^1 модели

$$f(a^1, b^1|Y_ч, Z) \sim \{S^2 + (a^1 - \hat{a}^1, b^1 - \hat{b}^1)ZZ^т(a^1 - \hat{a^1}, b^1 - \hat{b}^1)^т\}^{-\frac{N}{2}}, \qquad (14)$$

где $S^2 \triangleq (Y_ч - (\hat{a}^1, \hat{b}^1)Z)(Y_ч - (\hat{a}^1, \hat{b}^1)Z)^т \equiv \hat{\eta}\hat{\eta}^т$.

В подразделе 3.5.3. получены маргинальные функций плотности распределения вероятностей параметров a^1 и b^1 модели. Для параметра a^1 установлено, что

$$\begin{aligned} &f(a^1|Y_ч, Y_{нч}, X_{нч}) \sim \\ &\sim \{S^2 + (\delta_a)^2(Y_{нч}Y_{нч}^т - Y_{нч}X_{нч}^т(X_{нч}X_{нч}^т)^{-1}X_{нч}Y_{нч}^т)\}^{-\frac{N-1}{2}}, \end{aligned} \qquad (15)$$

где $(\delta_a) = (a^1 - \hat{a}^1)$.

Маргинальная апостериорная функция плотности распределения вероятностей параметра b^1 имеет вид

$$f(b^1|Y_ч, Y_{нч}, X_{нч}) \sim \\ \sim \{S^2 + (\delta_b)^2(X_{нч}X_{нч}^{\mathrm{T}} - X_{нч}Y_{нч}^{\mathrm{T}}(Y_{нч}Y_{нч}^{\mathrm{T}})^{-1}Y_{нч}X_{нч}^{\mathrm{T}})\}^{-\frac{N-1}{2}}, \tag{16}$$

где $(\delta_b) = (b^1 - \hat{b}^1)$, $S^2 \triangleq (Y_ч - (\hat{a}^1, \hat{b}^1)Z)(Y_ч - (\hat{a}^1, \hat{b}^1)Z)^{\mathrm{T}} \equiv \hat{\eta}\hat{\eta}^{\mathrm{T}}$.

Функция плотности распределения вероятностей параметра a^0, совместная функция плотности распределения вероятностей параметров a^1 и b^1 модели и маргинальные апостериорные функции плотности распределения вероятностей параметров a^1 и b^1 являются обобщенными функциями плотности распределения вероятностей Стьюдента.

В подразделе 3.5.4. описана методика построения доверительных интервалов для параметров модели с использованием найденной функции плотности распределения вероятностей параметра a^0 и маргинальных плотностей распределения вероятностей параметров a^1 и b^1 модели.

В разделе 3.6 описан метод определения вероятностей реализаций возможных сценариев изменения численностей взаимодействующих популяций нормальных и аномальных клеток с использованием K-кратной имитации значений $\{a_j^0, a_j^1\}_{j=1}^K$ параметров a^0 и a^1 с использованием байесовских апостериорных законов распределения, определения соответствующего сценария изменения численностей изучаемых популяций для каждой пары значений a_j^0, a_j^1 с использованием аналитического решения задачи Коши для линейной системы и использования комбинаторного определения вероятности.

Приведен пример получения оценки вероятностей реализации различных сценариев развития клеточной популяционной системы.

В **четвертой главе** решена задача параметрической идентификации нелинейной модели (6), предложенной в главе 2, по результатам экспериментов. Получен дискретный аналог нелинейной модели (6):

$$X_{k+1} = h_{01}^d X_k + h_{02}^d X_k^2 + h_{03}^d X_k Y_k, \\ Y_{k+1} = h_{11}^d X_k + h_{12}^d Y_k + h_{13}^d X_k^2 + h_{14}^d X_k Y_k + h_{15}^d Y_k^2, \tag{17}$$

где

$$h_{01}^d = \tilde{h}_{01}\Delta t + 1, \quad h_{02}^d = \tilde{h}_{02}\Delta t, \quad h_{03}^d = \tilde{h}_{03}\Delta t, \quad h_{11}^d = \tilde{h}_{11}\Delta t, \\ h_{12}^d = \tilde{h}_{12}\Delta t + 1, \quad h_{13}^d = \tilde{h}_{13}\Delta t, \quad h_{14}^d = \tilde{h}_{14}\Delta t, \quad h_{15}^d = \tilde{h}_{15}\Delta t, \\ \tilde{h}_{ij} = \frac{h_{ij}X_0}{\tau^0}, \tag{18}$$

а параметры h_{ij} заданы соотношениями (7).

Анализируемая модель является линейной относительно искомых параметров, что позволяет обобщить на исследуемый случай результаты, полученные в главе 3 для линейной модели. Предполагается, что измерения численностей проводятся через равные промежутки времени Δt.

В разделе 4.1 решена задача нахождения точечных оценок параметров нелинейной модели с учетом присущих ей особенностей. Показано, что задачу идентификации векторов параметров первого и вектора параметров второго уравнений модели (17) можно разбить на две независимые подзадачи: нахождение оценки параметров h_{01}^d, h_{01}^d, h_{03}^d, определяющих динамику развития популяции нормальных клеток, и оценок параметров $h_{11}^d, \ldots, h_{15}^d$, определяющих динамику развития популяции аномальных клеток.

В разделе 4.2 изложен алгоритм проверки статистической гипотезы о виде закона распределения случайных возмущений исходной модели состояния.

В разделе 4.3 на основе байесовского подхода и теории инвариантности Джеффриса построены апостериорные функции плотности распределения вероятностей для параметров нелинейной модели. Показано, что указанные функции подчиняются обобщенному закону Стьюдента.

Например, плотность совместного распределения вероятностей вектора параметров модели $H_0^d = (h_{01}^d, h_{02}^d. h_{03}^d)$ может быть представлена в следующем виде:

$$f(H_0^d | X_ч, Z_x) \sim \{S_0^2 + (H_0^d - \widehat{H}_0^d)(Z_x Z_x^{\intercal})(H_0^d - \widehat{H}_0^d)^{\intercal}\}^{-\frac{N}{2}}. \qquad (19)$$

Здесь $X_ч \triangleq (X_2, \ldots, X_{2N}) \in M_{1\times N}(R)$, $X_{нч} \triangleq (X_1, \ldots, X_{2N-1}) \in M_{1\times N}(R)$, а $Z_x = \left(X_{нч}^{\intercal}, X_{нч}^{2\,\intercal}, (X_{нч}Y_{нч})^{\intercal}\right)^{\intercal}$, $\widehat{H}_0^d = X_ч Z_x^+$ — оценка вектора параметров H_0^d, а $S_0^2 \triangleq (X_ч - \widehat{H}_0^d Z_x)(X_ч - \widehat{H}_0^d Z_x)^{\intercal}$.

В разделе 4.4 найдены маргинальные апостериорные функции плотности распределения вероятностей всех параметров математической модели (17). Например, маргинальная апостериорная функция плотности распределения вероятностей параметра h_{02}^d может быть представлена как

$$f(h_{02}^d | X_ч, Z_x) \sim \\ \sim \{S_0^2 + (h_{02}^d - \hat{h}_{02}^d)^2 (Z_{1x}Z_{1x}^{\intercal} - Z_{1x}Z_{2x}^{\intercal}(Z_{2x}Z_{2x}^{\intercal})^{-1}Z_{2x}Z_{1x}^{\intercal})\}^{-\frac{N-2}{2}}, \qquad (20)$$

где $S_0^2 = (X_ч - \widehat{H}_0^d Z_x)(X_ч - \widehat{H}_0^d Z_x)^{\intercal}$, а $Z_{1x} = X_{нч}^2$ и $Z_{2x} = (X_{нч}^{\intercal}; (X_{нч}Y_{нч})^{\intercal})^{\intercal}$.

В разделе 4.5 описан алгоритм построения доверительных интервалов для параметров модели на основе полученных маргинальных апостериорных функции плотности распределения вероятностей.

Основные результаты диссертационной работы

1. Разработана и исследована линейная модель развития клеточных популяций нормальных и аномальных клеток при культивировании в стандартных лабораторных условиях, отражающая поведение клеточных популяций при отсутствии ограничений на ресурсы.

2. Разработана и исследована нелинейная модель, предназначенная для моделирования динамики развития клеточной популяционной системы в условиях ограниченных ресурсов.

3. На основе баесовского подхода при ограниченных выборках экспериментальных данных получены совместные апостериорные функции плотности распределения вероятностей параметров предложенных математических моделей, а также маргинальные плотности распределения вероятностей этих параметров.

4. Получены интервальные оценки для параметров моделей с использованием найденных маргинальных функций плотности распределения вероятностей.

5. Разработана методика использования точечных оценок параметров линейной модели и законов распределения параметров для анализа возможных сценариев развития изолированной популяционной системы.

Основные результаты диссертации отражены в работах

1. Статистический анализ клонообразования в культурах стволовых клеток человека/ М.С. Виноградова [и др.] // Клеточные технологии в биологии и медицине. 2011. № 2. С. 63–66.

2. Математическая модель суммарных численостей взаимодействующих клеточных популяций / М.С. Виноградова [и др.] // Вестник Моск. гос. техн. ун-та им. Н.Э. Баумана. Сер.: Естественные науки. 2011. № 1. С. 18–24.

3. Бочков Н.П., Виноградова М.С., Волков И.К. Оценка вероятности реализации вариантов развития взаимодействующих клеточных популяций // Вестник Моск. гос. техн. ун-та им. Н.Э. Баумана. Сер.: Естественные науки. 2011, № 3. С. 31–43.

4. Виноградова М.С. Качественный анализ модели функционирования взаимодействующих клеточных популяций // Наука и образование. Электрон. журнал. МГТУ им. Н.Э. Баумана. 2011. № 11. С. 1-20. Режим доступа: http://technomag.edu.ru/doc/251409.html.

5. Виноградова М.С. Параметрическая идентификация модели взаимодействующих клеточных популяций на основе байсовского подхода // Наука и образование. Электрон. журнал. МГТУ им. Н.Э. Баумана. 2012. № 11. С. 155–182. Режим доступа: http://technomag.edu.ru/doc/490900.html.

6. Виноградова М.С. Построение маргинальных плотностей распределения вероятности параметров модели клеточной популяции// Наука и образование. Электрон. журнал. МГТУ им. Н.Э. Баумана. 2012. № 12. С. 121–134. Режим доступа: http://technomag.edu.ru/doc/500575.html.

7. Виноградова М.С. Математическая модель клеточной популяции // Математика. Компьютер. Образование: тез. докл. XVIII Международной конф. Пущино, 2011. С. 25.

8. Статистическое обоснование критериев клонообразования в культурах стволовых клеток человека / М.С. Виноградова [и др.] // Молекулярная генетика соматических клеток: тез. докл. VIII Междунар. конф. Звенигород, 2011. С. 79–80.