

УДК 517.958:57

*М. Н. Подлевских, З. В. Шилова*

## МЕТОДЫ И МОДЕЛИ ДИСКРЕТНОЙ МАТЕМАТИКИ В БИОЛОГИИ

В данной статье приведен обзор основных классификаций математических моделей, которые могут быть использованы в биологии. Отбор материала обусловлен необходимостью методического обоснования его изучения студентами биологических направлений подготовки. Интерес к приложениям дискретной математики обусловлен не только развитием биоинформатики и математической биологии, но и необходимостью корректировки содержания математической составляющей в профессиональной подготовке специалистов естественно-научных профилей с учетом появления и внедрения новых профессиональных стандартов. Авторами рассмотрены базовые понятия дискретной математики (множества, бинарные отношения, графы), описаны простые математические модели биологического содержания, использующие теории марковских и ветвящихся процессов, указаны некоторые математические методы, используемые в биологических исследованиях, а также в смежных с ней отраслях знаний – экологии и медицине.

*Ключевые слова:* математическая модель в биологии, дискретная математика, методы дискретной математики, цепи Маркова.

Использование математики в исследовании предметной области реализуется в виде такого научного метода как математическое моделирование. Существующие учебные курсы по изучению математических методов и моделей в биологии и экологии посвящены в основном моделям, инструментом для создания и анализа которых выступает аппарат непрерывной математики. Результаты, полученные в этой области, сыграли огромную роль в становлении нового научного направления, которое получило название математической биологии. На рубеже XX и XXI веков это научное направление получило новый импульс развития благодаря использованию современных информационных

технологий. Поэтому в настоящее время говорят о такой области знаний как математическая биология и биоинформатика, в которой биологические объекты изучаются методами математики и информатики. При этом очевидно, что значительно расширяется круг используемых математических понятий, теорий, методов. Поэтому и в подготовке современных специалистов в области биологии и смежных с ней областей знаний назревает необходимость корректировки содержания их математического образования. Наряду с изучением классических разделов высшей математики требуется первоначальное знакомство с элементами дискретной математики. Отметим, что традиционные вопросы дискретной математики присутствуют, но весьма дозировано, в программах дисциплин, посвященных информатике и информационным технологиям, где, в свою очередь, не уделяется внимание формированию у обучающихся представлений о математических моделях.

Анализ научной литературы [1, 2, 3-5, 6, 7, 8] показывает, что решение задач описания биологических, медицинских или экологических объектов, на сегодняшний день, предполагает достаточный уровень абстрагирования, сравнимый с используемым в математических теориях. В создании математических моделей биологических объектов и процессов, в том числе возникающих в смежных областях знаний в экологии и медицине, все больше находят применение методы абстрактной алгебры, теории графов, теории конечных автоматов, дискретной математики и другие. Например, в работе [7] И. И. Слеповичева, посвященной искусственным нейронным сетям, даются формулировки и определения нейросетевых моделей, изучаются их свойства и возможности с точки зрения теории графов и универсальной алгебры.

Применение математики в биологии теснейшим образом связано в том числе и с задачами обработки результатов наблюдений и экспериментов. Фундаментальное методологическое понятие «математическая модель» позволяет объединить эти две составляющие математического обеспечения биологических исследований.

Существуют различные классификации математических моделей, рассмотрим основные из них. Так, например, по *способу построения* все модели делят на два класса: материальные и абстрактные. Материальные модели по своей физической природе сходны с оригиналом. Они могут сохранять геометрическое подобие оригиналу (макеты, тренажеры, искусственные заменители органов и т. д.), подобие протекания физических процессов с оригиналом – физическое моделирование (гидродинамическая модель – течение воды и т. п.) и могут быть природными объектами – прообразами оригинала, то есть натурными моделями (метод пробных участков). Абстрактные модели, представляющие собой описание оригинала в словесной форме или посредством символов и операций над ними, отражающих исследуемые особенности оригинала, подразделяются на три типа: вербальные, схематические и математические. Вербальные модели – это формализованный вариант традиционного естественнонаучного описания в виде текста, таблиц и иллюстраций [9]. Схематические модели разрабатываются в виде различного рода схем, рисунков, графиков и фотографий, основные их достоинства – наглядность, информативность и простота построения (трофические цепи, пирамида Элтона, схемы структуры, динамики и энергетики экосистем, воздействия экологических факторов, биохимических круговоротов, и др.). Вербальные и схематические модели – неотъемлемая часть качественного анализа математического моделирования, являющегося наиболее совершенным видом количественного исследования оригинала, позволяющая построить его математическую модель, (Лиела, 1982 г.) практически такая модель представляет собой формулу или систему уравнений и неравенств.

По своему *характеру* выделяют модели статические и динамические. Статическая модель отражает объект (систему), не изменяющий свое состояние во времени, а динамическая модель отражает объект (систему), изменяющий свое состояние во времени. Отметим, что подавляющее большинство живых

объектов и систем – это динамические системы и могут быть отражены только лишь динамическими моделями.

В соответствии с объектом и целями математические модели в биологии можно подразделить на три больших класса. Первый класс – регрессионные модели, включает эмпирически установленные зависимости (формулы, дифференциальные и разностные уравнения, статистические законы), не претендующие на раскрытие механизма изучаемого процесса. Второй класс – это имитационные модели конкретных сложных живых систем, как правило, максимально учитывающие имеющуюся информацию об объекте. Имитационные модели применяются для описания объектов различного уровня организации живой материи как от биомакромолекул, так и до моделей биогеоценозов. Третий класс – качественные (базовые) модели – простые модели, которые поддаются аналитическому исследованию и обладают свойствами, которые позволяют описывать целый спектр природных явлений. Все биологические системы различного уровня организации, начиная от биомакромолекул и вплоть до популяций, являются термодинамически неравновесными, открытыми для потоков вещества и энергии. Поэтому нелинейные динамические модели выступают в роли базовых моделей математической биологии.

Классификация, для которой основным средством построения математических моделей является статистика: модели данных и модели систем. При построении *модели данных* основной целью моделирования является поиск математической функции, наиболее точно описывающей имеющийся набор данных, полученных экспериментально. *Модели систем* основаны на конкретных гипотезах о структуре и принципах функционирования изучаемого явления, такие модели предназначены для теоретического изучения механизмов явления или структуры системы. В этом случае для построения теоретических моделей часто используются методы дискретной математики: алгебраические, теоретико-множественные, логические, сетевые, графовые и т. п.

Принципы моделирования биологических систем проанализированы П. М. Брусиловским в [2], где он дает характеристику математической биологии как одного из направлений в изучении феномена жизни. В этой же статье отмечается расширение сферы применения метода математического моделирования при исследовании функционирования биосистем, которое в основном происходит по двум направлениям, соответствующим экстенсивному (увеличение числа моделируемых биосистем) и интенсивному (увеличение «глубины» изучения биосистем) развитию научного познания. Группа авторов, В. К. Шитиков, Г. С. Розенберг, Т. Д. Зинченко, в своей работе [10] приводят расширенный обзор методов математического моделирования применительно к экологическому мониторингу пресноводных водоемов. В то же время анализ математических методов исследования проводится с позиций, позволяющих дать общие представления об особенностях моделирования биосистем. Большое число математических моделей дискретного характера, в том числе биологического содержания, содержится в книге Ф. Робертса [6]. Классическим научным изданием, содержащим подробное изложение разнообразных математических моделей в биологии и медицине, является книга Н. Бейли [1].

Рассмотрим далее примеры базовых понятий, моделей и методов дискретной математики, которые могут применяться в биологии, а также в смежных с ней отраслях знаний: экологии и медицине. В силу того, что отбор материала обусловлен необходимостью методического обоснования его изучения студентами биологических направлений подготовки, в данной статье он представлен в совокупности трех уровней.

Во-первых, это выделение некоторых базовых **понятий** дискретной математики и краткие иллюстрации их использования в биологии. Во-вторых, демонстрация конкретных математических **моделей** дискретного характера, а также в обзорном плане, учитывая объем статьи, перечисление связей методов дискретной математики и смежных с ней информационных и математических теорий, используемых в биологии.

Под дискретной математикой понимают область математических знаний обо всех видах дискретных объектов из теории чисел, алгебры, математической логики, геометрии и других разделов математики. В качестве примеров базовых понятий дискретной математики выделим: множества, бинарные отношения (отношение эквивалентности), графы.

*Множества и операции над ними.* В естествознании важным инструментом описания биоразнообразия является классификация, для проведения которой вводятся различные приемы сравнения объектов на основе использования мер сходства и различия. Большинство выражений для индексов сходства основаны на общих положениях теории множеств, которые могут быть интерпретированы в виде диаграммы Эйлера – Венна (рис. 1). При использовании конкретных выражений для коэффициентов подобия в формулы могут подставляться мощности (число элементов) подмножеств  $a$ ,  $b$ ,  $c$  и  $d$ , если исследователи хотят ограничиться альтернативными высказываниями «отсутствие / наличие» вида, либо показатели обилия в абсолютной или интервальной шкале. В первом случае мы будем отождествлять мощность подмножества с ним самим.



Рис. 1.

Первая попытка количественного выражения степени сходства между сообществами принадлежала в 1901 г. швейцарскому исследователю П. Жаккару. Коэффициент флористического сходства Жаккара до сих пор широко используется в геоботанике:

$$K_j = \frac{c}{a+b+c}.$$

Понятие *дескриптивного множества* используется в теории измерений, являющейся базой экспериментального метода исследования – основного метода в естественных науках. Дескриптивным множеством называется конечное множество, каждому элементу которого поставлено в соответствие неотрицательное число («вес»).

*Отношение эквивалентности* математически определяет понятия «тождественности», «схожести» и играет большую роль при образовании новых абстрактных объектов. Меры сходства (близости) естественно считаются эквивалентными, если результаты оценки близости при замене одной меры другой, останутся качественно теми же: более близкие объекты останутся более близкими, а более удаленные – более удаленными, то есть при использовании эквивалентных мер близости в задачах классификации должны получаться одни и те же результаты. Большое внимание строгому определению эквивалентности мер сходства и различия было уделено Б. И. Сёмкиным [11].

При изучении структуры биоразнообразия, например, в экологической системе, приходится выявлять различные бинарные отношения: отношения принадлежности, отношения сравнения, отношения различия и другие. Естественным математическим аппаратом при этом является использование *графов*. Простые примеры использования графов в дискретных математических моделях биологического содержания можно найти, например, в работе А. В. Галанина [12].

В качестве примера дискретной математической модели биологического содержания можно привести модель, описывающую процесс непрерывного скрещивания с гибридом, построенную на использовании *цепи Маркова*. Для построения такой модели нужно вспомнить некоторые факты из теории наследственности и учитывать вероятностный характер результатов. Согласно базовым положениям теории наследственности, появление у потомства признаков строго определяется набором генов, которые случайным образом получены от родителя. Согласно Менделю, при скрещивании двух особей

потомство получает по одному гену от каждого родителя. Выбор гена происходит случайным образом и независимо у каждого из родителей. Если скрещиваются две чисто доминантные  $dd$ -особи по какому-то из признаков, потомок получает по  $d$ -гену от каждого родителя и, следовательно, будет иметь тип  $dd$ , то есть также чисто доминантным по данному признаку. При скрещивании чисто доминантной  $dd$ -особи с чисто рецессивной  $rr$ -особью потомок получает один ген от своего чисто доминантного родителя и другой ген от чисто рецессивного родителя; следовательно, он является гибридом, но по внешнему виду его нельзя отличить от чистого доминанта. Если скрещиваются чистый доминант  $dd$  и гибрид  $dr$ , потомок получает один ген  $d$  от  $dd$ -родителя, а другой ген  $d$  или  $r$  от  $dr$ -родителя. Последние две возможности имеют одинаковую вероятность. Поэтому потомок является чистым доминантом с вероятностью  $1/2$  и гибридом с вероятностью  $1/2$ . Пусть гибрид  $dr$  скрещивается с другим гибридом  $dr$ . Потомок получает с вероятностью  $1/2$  ген  $d$  или  $r$  от первого родителя и так же от второго родителя. Следовательно, потомок оказывается типа  $dd$ ,  $rd$  (что генетически равноценно  $dr$ ) или  $rr$ . Вероятность получения потомка чистого доминанта  $dd$  равна  $1/4$ , гибрида  $dr$  или  $rd$  равна  $1/2$ , а вероятность получения чистой рецессивной особи  $rr$  равна  $1/4$ . Следовательно, если скрещиваются два гибрида, то с вероятностью  $1/4$  потомок будет иметь рецессивные признаки, а с вероятностью  $3/4$  – доминантные.

На основе этой модели можно получить и другие результаты существенно более сложные. Например, особь неизвестного генетического происхождения скрещивается с известным гибридом. Случайным образом выбирается один из потомков и снова скрещивается с известным гибридом, и т. д. Анализ процесса непрерывного скрещивания с гибридом может быть получен с помощью цепей Маркова. В марковской цепи информация о предшествующих испытаниях, а также относительно номера испытания не влияет на вероятности будущих событий, если известен последний исход.

Построим цепь Маркова, в которой состояниями являются генотипы особи, выбранной из  $t$ -го поколения. Следовательно, имеются три состояния:  $D$  ( $i, j = 1$ ) – «чистая доминантная особь»;  $H$  ( $i, j = 2$ ) – «гибрид»;  $R$  ( $i, j = 3$ ) – «чистая рецессивная особь». Переходные вероятности для этой цепи задаются следующей матрицей  $P$  (рис. 2):

	$D$	$H$	$R$
$D$	1/2	1/2	0
$H$	1/4	1/2	1/2
$R$	0	1/2	1/2

Рис. 2.

Элемент  $p(i, j)$  матрицы  $P$  определяет вероятность того, что, если особь генетического типа  $i$  ( $i = 1, 2, 3$ ) скрещивается с известным гибридом, то потомок, выбранный случайным образом, будет иметь генетический тип  $j$  ( $j = 1, 2, 3$ ), так, например,  $p(1, 1) = 1/2$ , потому что при скрещивании чистого доминанта с гибридом вероятность получения потомка – чистого доминанта – равна  $1/2$ . Переходный орграф, соответствующий матрице  $P$ , показан на рис. 3.

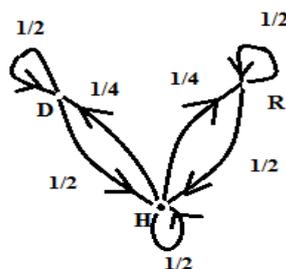


Рис. 3.

Данный граф сильно связан, поэтому цепь Маркова является эргодической. Наличие петель определяет регулярность цепи. Стационарный вероятностный вектор матрицы  $P$  имеет вид  $(1/4, 1/2, 1/4)$ , что означает следующее: при длительном повторении процесса скрещивания независимо от генотипа особи, с которой процесс скрещивания начался, в  $t$ -м поколении

произвольно взятая особь будет чисто доминантной с вероятностью  $1/4$ , гибридной с вероятностью  $1/2$  и чисто рецессивной с вероятностью  $1/4$ .

Возникновение и развитие математических теорий в большинстве случаев связано с необходимостью решать практические задачи из разных областей знаний. История применения математики в биологии, медицине и других смежных с ними наук приводит немало подобных примеров. Широко известным является факт создания и развития статистических методов анализа, так, пытаясь математически оформить теорию наследственности Гальтона, Пирсон в 1898 г. разработал основы множественной регрессии, а позже основы теории сопряжения признаков и основы нелинейной корреляции и регрессии.

Другим интересным примером создания новой математической теории из решения практических задач является теория **ветвящихся процессов**. Она возникла из попыток объяснить причины вырождения знаменитых фамилий в Великобритании. В настоящее время эта теория, как раздел теории вероятностей, имеет широкие приложения в теории алгоритмов, теории массового обслуживания, а также в приложениях физического, химического и биологического характера.

Дискретная модель ветвящегося процесса Гальтона – Ватсона позволяет, например, найти вероятностные оценки для продолжительности периода сохранения или вырождения некоторой популяции. В работе В.А. Ватутина [13] приводится пример для популяции североатлантических китов. Самка этих животных может произвести за год либо ни одной, либо одну, либо две особи женского пола. При этом смерть матери влечет гибель потомства за этот год. При построении модели в момент времени  $n$  (промежуток длиною в год) принимаются следующие условия: самка не производит потомства, если она погибает до начала  $(n + 1)$ -го года; производит одну особь женского пола (себя), если она выживает, но не производит особей женского пола; производит две особи женского пола (себя и одного потомка), если она не погибает за  $n$ -й год; других возможностей нет. Если за  $p$  обозначается вероятность дожития

самки до следующего года, а за  $q$  – вероятность рождения потомка женского пола, то вероятностная производящая функция согласно условиям, указанным выше, имеет следующий вид:  $f(s) = 1 - p + p(1 - q)s + pqs^2$ . При этом среднее равно . Для значений используемых в работе [13] параметров:  $p = 0,94$ ;  $q = 0,051$  получается, что среднее  $A$  равно  $0,988$  . В результате оценка снизу для числа поколений (лет)  $n$ , по истечении которых популяция китов сохранится с вероятностью, большей  $0,99$  имеет вид  $n \geq 357$  . Оценка сверху для количества лет, по истечении которых популяция исчезнет с вероятностью, большей  $0,99$  составляет  $n \leq 789$  .

Во многих учебных планах подготовки биологов уровня бакалавриата и, особенно, магистратуры имеются курсы, в которых происходит знакомство с применением математических методов и моделей в биологии. Считаем, что содержание таких учебных дисциплин будет более разнообразным, если обучающиеся смогут использовать при знакомстве с ними методы дискретной математики [14, 15].

### Список литературы

1. Бейли Н. Математика в биологии и медицине. М.: Мир, 1970. 326 с.
2. Брусиловский П. М. Становление теоретической биологии и математическое моделирование // Проблемы анализа биологических систем. М.: Изд-во МГУ, 1983.
3. Дромашко С. Е. Математическое и компьютерное моделирование в биологии: учеб.-метод. пособие. Минск: ИПНК, 2009. 65 с.
4. Евдокимов А. А. Математические методы анализа дискретных моделей генных сетей // Новые информационные технологии в исследовании сложных структур: тез. докл. IX конф. Томск, 2012. С. 73.
5. Журавлев Ю. И. Петров И. Б., Рязанов В. В. Дискретные методы диагностики и анализа медицинской информации // Медицина в зеркале информатики: сб. РАН. М.: Наука, 2008. С. 113–123.
6. Робертс Ф. С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. М.: Наука, 1986. 496 с.

7. Слеповичев И. И. Алгебраические свойства абстрактных нейронных сетей. // Известия Саратовского университета. Сер. «Математика. Механика. Информатика» 2016. Т. 16. Вып. 1. С. 96–103.
8. Ризниченко Г. Ю., Рубин А. Б. Математические модели биологических продукционных процессов: учеб. пособие. М.: Изд-во МГУ, 1993. 302 с.
9. Федоров В. Д., Гильманов Т. Г. Экология. М.: Изд-во МГУ, 1980. 464 с.
10. Шитиков В. К., Розенберг Г. С., Зинченко Т. Д. Количественная гидроэкология: методы системной идентификации. Тольятти: ИЭВБ РАН, 2003. 463 с.
11. Семкин Б. И., Горшков М. В. Аксиоматическое введение мер сходства, различия, совместимости и зависимости для компонентов биоразнообразия // Вестник ТГЭУ. 2008. № 4. С. 31–46.
12. Галанин А. В. Флора и ландшафтно-экологическая структура растительного покрова. Владивосток: ДВО АН СССР, 1991. 272 с.
13. Ватулин В. А. Лекционные курсы НОЦ / Математический институт им. В. А. Стеклова РАН (МИАН). Вып. 8. М.: МИАН, 2008. 108 с.
14. Подлевских М. Н. Интерактивный ресурс по математике для студентов биологических и экологических направлений подготовки // Вестник ВятГГУ. Сер. «Педагогика». 2015. № 4. С. 116–120.
15. Подлевских М. Н. Использование графов в дискретных математических моделях биологического и экологического содержания // Математический вестник педвузов и университетов Волго-Вятского региона: межвуз. сб. науч.-метод. работ. Вып. 17. Киров: Изд-во ООО «Радуга-ПРЕСС», 2015. С. 145–150.

**ПОДЛЕВСКИХ Марина Николаевна** – доцент кафедры математического моделирования, Вятский государственный университет. 610000, г. Киров, ул. Московская, 36.

E-mail: kspmar@yandex.ru

**ШИЛОВА Зоя Вениаминовна** – доцент кафедры фундаментальной и компьютерной математики, Вятский государственный университет. 610000, г. Киров, ул. Московская, 36.

E-mail: zoya@soi.su