

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
"Казанский (Приволжский) федеральный университет"  
Институт вычислительной математики и информационных технологий



**УТВЕРЖДАЮ**

Проректор  
по образовательной деятельности КФУ  
Проф. Таюрский Д.А.

\_\_\_\_\_ 20\_\_ г.

**Программа дисциплины**

Вероятностные модели генетики Б1.В.ДВ.13

Направление подготовки: 01.03.04 - Прикладная математика

Профиль подготовки: Математическое моделирование

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

**Автор(ы):**

Володин И.Н. , Салимов Р.Ф. , Симушкин С.В.

**Рецензент(ы):**

Халиуллин С.Г.

**СОГЛАСОВАНО:**

Заведующий(ая) кафедрой: Турилова Е. А.

Протокол заседания кафедры No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 201\_\_ г

Учебно-методическая комиссия Института вычислительной математики и информационных технологий:

Протокол заседания УМК No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 201\_\_ г

Регистрационный No

## Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) профессор, д.н. (профессор) Володин И.Н. кафедра математической статистики отделение прикладной математики и информатики , igornvolodin@gmail.com ; старший преподаватель, б/с Салимов Р.Ф. кафедра математической статистики отделение прикладной математики и информатики , Rustem.Salimov@kpfu.ru ; доцент, к.н. (доцент) Симушкин С.В. кафедра математической статистики отделение прикладной математики и информатики , Sergey.Simushkin@kpfu.ru

### 1. Цели освоения дисциплины

Дисциплина принадлежит к вариативной части подготовки бакалавров по направлению "Прикладная математика". Цель освоения дисциплины - изучение современных компьютерных технологий в области математических вычислений и приобретение навыков применения специализированных математических пакетов в научной деятельности.

### 2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел "Б1.В.ДВ.13 Дисциплины (модули)" основной образовательной программы 01.03.04 Прикладная математика и относится к дисциплинам по выбору. Осваивается на 4 курсе, 8 семестр.

Курс рассчитан на бакалавров, имеющих подготовку по дисциплинам "Математический анализ", "Алгебра и геометрия", "Дифференциальные уравнения", "Уравнения математической физики", "Информатика". Предполагается, что студенты знакомы с основами математического моделирования и дифференциальными уравнениями, описывающими физические процессы. Знания, навыки и умения, приобретенные в результате прохождения курса, будут востребованы при изучении специальных курсов, касающихся сложных математических вычислений с применением ЭВМ, а также при выполнении научных работ, необходимых для получения квалификации.

### 3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-2 (профессиональные компетенции)	способность использовать современные математические методы и современные прикладные программные средства и осваивать современные технологии программирования
ПК-10 (профессиональные компетенции)	готовность применять математический аппарат для решения поставленных задач, способность применить соответствующую процессу математическую модель и проверить ее адекватность, провести анализ результатов моделирования, принять решение на основе полученных результатов
ПК-12 (профессиональные компетенции)	способность самостоятельно изучать новые разделы фундаментальных наук
ПК-9 (профессиональные компетенции)	способность выявить естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, готовность использовать для их решения соответствующий естественнонаучный аппарат

В результате освоения дисциплины студент:

1. должен знать:

основные вероятностные модели генетики;

2. должен уметь:

строить вероятностную модель, оценивать параметры модели

3. должен владеть:

навыками выбора методов построения вероятностных моделей;

4. должен демонстрировать способность и готовность:

способность и готовность применять на практике полученные знания в области вероятностных и статистических методов математической генетики

#### 4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) 72 часа(ов).

Форма промежуточного контроля дисциплины: зачет в 8 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);

55-70 баллов - "удовлетворительно" (удов.);

54 балла и менее - "неудовлетворительно" (неуд.).

#### 4.1 Структура и содержание аудиторной работы по дисциплине/ модулю Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практи- ческие занятия	Лабора- торные работы	
1.	Тема 1. Анализ одной последовательности ДНК	8	1-3	0	0	8	Письменное домашнее задание
2.	Тема 2. Анализ нескольких последовательностей ДНК или последовательностей протеинов	8	4-7	0	0	8	Письменное домашнее задание
3.	Тема 3. Базовый метод поиска локальных выравниваний	8	8-10	0	0	7	Письменное домашнее задание
4.	Тема 4. Экспрессия генов, микрочипы и многомерный анализ	8	11-13	0	0	7	Письменное домашнее задание

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
5.	Тема 5. Модели эволюции	8	14-16	0	0	6	Письменное домашнее задание
6.	Тема 6. Оценки на филогенетических деревьях	8	16-18	0	0	4	Контрольная работа
.	Тема . Итоговая форма контроля	8		0	0	0	Зачет
	Итого			0	0	40	

#### 4.2 Содержание дисциплины

##### Тема 1. Анализ одной последовательности ДНК

###### лабораторная работа (8 часа(ов)):

1. Краткое введение в генетику 2. Построение простейшей модели ДНК 3. Анализ шаблонов 4. Перекрытие последовательностей

##### Тема 2. Анализ нескольких последовательностей ДНК или последовательностей протеинов

###### лабораторная работа (8 часа(ов)):

1. Частотное сравнение 2. Простой тест схожести 3. Алгоритмы выравнивания двух последовательностей 4. Последовательности протеинов и матрицы переходов

##### Тема 3. Базовый метод поиска локальных выравниваний

###### лабораторная работа (7 часа(ов)):

1. Введение в BLAST 2. Сравнение двух выравненных последовательностей 3. Сравнение двух невыравненных последовательностей

##### Тема 4. Экспрессия генов, микрочипы и многомерный анализ

###### лабораторная работа (7 часа(ов)):

1. Статистический анализ данных с микрочипов 2. Дисперсионный анализ 3. Главные компоненты

##### Тема 5. Модели эволюции

###### лабораторная работа (6 часа(ов)):

1. Модели переходов нуклеотидов 2. Модели дискретного времени 3. Модели непрерывного времени

##### Тема 6. Оценки на филогенетических деревьях

###### лабораторная работа (4 часа(ов)):

1. Расстояние на деревьях 2. Предполагаемые расстояния 3. Максимальное правдоподобие

#### 4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел дисциплины	Се-местр	Неде-ля семе-стра	Виды самостоятельной работы студентов	Трудо-емкость (в часах)	Формы контроля самостоятельной работы
	Тема 1. Анализ					

одной последовательности ДНК

---

подготовка домашнего задания

письмен-  
ное

домаш-  
нее задание



№	Раздел дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
2.	Тема 2. Анализ нескольких последовательностей ДНК или последовательностей протеинов	8	4-7	подготовка домашнего задания	6	письменное домашнее задание
3.	Тема 3. Базовый метод поиска локальных выравниваний	8	8-10	подготовка домашнего задания	5	письменное домашнее задание
4.	Тема 4. Экспрессия генов, микрочипы и многомерный анализ	8	11-13	подготовка домашнего задания	5	письменное домашнее задание
5.	Тема 5. Модели эволюции	8	14-16	подготовка домашнего задания	5	письменное домашнее задание
6.	Тема 6. Оценки на филогенетических деревьях	8	16-18	подготовка к контрольной работе	5	контрольная работа
	Итого				32	

### 5. Образовательные технологии, включая интерактивные формы обучения

Предполагается использование диалоговой формы ведения занятий с постановкой и решением проблемных задач, обсуждением дискуссионных моментов.

Некоторые разделы курса студенты изучают самостоятельно по указанным методическим материалам или по Интернет-источникам с последующим докладом в виде презентаций и дискуссией.

При проведении практических занятий, на которых теоретический материал применяется к конкретным данным, студентам предлагается разработать самостоятельные вычислительные процедуры.

Перед каждым занятием, как лекционной, так и практической направленности проводится экспресс-опрос по пройденному теоретическому материалу.

### 6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

### **Тема 1. Анализ одной последовательности ДНК**

письменное домашнее задание , примерные вопросы:

Построение простой модели для последовательности ДНК

### **Тема 2. Анализ нескольких последовательностей ДНК или последовательностей протеинов**

письменное домашнее задание , примерные вопросы:

Сравнение двух данных последовательностей, с оценкой меры схожести.

### **Тема 3. Базовый метод поиска локальных выравниваний**

письменное домашнее задание , примерные вопросы:

Сравнение двух невыравненных последовательностей: многомерный анализ

### **Тема 4. Экспрессия генов, микрочипы и многомерный анализ**

письменное домашнее задание , примерные вопросы:

Применение дисперсионного анализа для сравнения экспрессии генов.

### **Тема 5. Модели эволюции**

письменное домашнее задание , примерные вопросы:

Расчёт матрицы переходных вероятностей для конкретной модели

### **Тема 6. Оценки на филогенетических деревьях**

контрольная работа , примерные вопросы:

1) Построение оценок по методу максимального правдоподобия. 2) Вычисление меры схожести двух последовательностей. 3) Выравнивание двух последовательностей.

### **Итоговая форма контроля**

зачет (в 8 семестре)

Примерные вопросы к итоговой форме контроля

На зачёте потребуются решение одной из следующих задач:

1. Используя совместную функцию плотности, докажите равенство 5.4
2. Покажите, что если  $L$  и  $G$  зафиксированы, среднее число контигов в 5.1 максимально (как функция от  $N$ ) при  $N=G/L$
3. Используя аппроксимацию (B.21), получите выражение для среднего значения размера контига 5.2, когда  $a$  мало. Поясните результат
4. Докажите выражение (5.26)
5. Для Днк длины 6 найдите вероятность, что слово  $gaga$  встретится 0, 1, и 2 раза.
6. Какие изменения нужны в формуле 5.44 для случаев  $N=4$  и 5.
7. Для случая равных вероятностей, найдите среднее число позиций до первого появления слова  $gagg$
8. Докажите выражение 5.93
9. Используя 5.99, найти среднее число позиций между успешными повторениями мотивов  $aaa$  и  $ata$
10. Подогнать  $x=cttgac$  к  $y=cagtatcgac$  с помощью скоринговой схемы из примера.

#### **7.1. Основная литература:**

1. Сазанов А.А. Генетика. - СПб.: ЛГУ им. А.С. Пушкина, 2011. - 264 с. Режим доступа: <http://znanium.com/bookread2.php?book=445036>
2. Бородин, А.Н. Элементарный курс теории вероятностей и математической статистики [Электронный ресурс] : учебное пособие / А.Н. Бородин. - Электрон. дан. - Санкт-Петербург : Лань, 2011. - 256 с. - Режим доступа: <https://e.lanbook.com/book/2026>

3. Боровков, А.А. Математическая статистика [Электронный ресурс] : учебник / А.А. Боровков. - Электрон. дан. - Санкт-Петербург : Лань, 2010. - 704 с. - Режим доступа: <https://e.lanbook.com/book/3810>
4. Лагутин, М.Б. Наглядная математическая статистика [Электронный ресурс] : учебное пособие / М.Б. Лагутин. - Электрон. дан. - Москва : Издательство 'Лаборатория знаний', 2015. - 475 с. - Режим доступа: <https://e.lanbook.com/book/70706>
5. Свешников, А.А. Прикладные методы теории вероятностей [Электронный ресурс] : учебник / А.А. Свешников. - Электрон. дан. - Санкт-Петербург : Лань, 2012. - 480 с. - Режим доступа: <https://e.lanbook.com/book/3184>

## 7.2. Дополнительная литература:

1. Сазанов А.А. Основы генетики. - СПб.: ЛГУ им. А.С. Пушкина, 2012. - 240 с. Режим доступа: <http://znanium.com/bookread2.php?book=445015>
2. Володин И. Н. Лекции по теории вероятностей и математической статистике [Текст: электронный ресурс] : [учебник] для студентов высших учебных заведений/ И. Н. Володин ; Казан. гос. ун-т, Каф. мат. статистики .? Электронные данные (1 файл: 1,5 Мб) .? (Казань : Казанский федеральный университет, 2013) .? Загл. с экрана .? Для 4-го и 5-го семестров .? Документ является электронной копией оригинала: Лекции по теории вероятностей и математической статистике: для студентов вузов/ И. Н. Володин. -- Казань: Казанский государственный университет, 2006. -- Режим доступа: открытый.  
URL:[http://libweb.kpfu.ru/ebooks/09\\_66%20\\_ds006.pdf](http://libweb.kpfu.ru/ebooks/09_66%20_ds006.pdf)
3. Симушкин С. В. Задачи по теории вероятностей [Текст: электронный ресурс] : учебное пособие / С. В. Симушкин, Л. Н. Пушкин .? Электронные данные (1 файл: 1,48 Мб) .? (Казань : Казанский федеральный университет, 2014) .? Загл. с экрана .? Режим доступа: открытый .  
Оригинал копии: Задачи по теории вероятностей : учебное пособие / С. В. Симушкин, Л. Н. Пушкин .? Казань : Казанский университет, 2011 .? 224 с.  
URL:<http://libweb.kpfu.ru/ebooks/publicat/0-787673.pdf>

## 7.3. Интернет-ресурсы:

- <http://learn.genetics.utah.edu/> - <http://learn.genetics.utah.edu/>  
<http://www.life.illinois.edu/ib/201/lectures/> - <http://www.life.illinois.edu/ib/201/lectures/>  
<http://www.stat.washington.edu/thompson/S394/genetics.html> -  
<http://www.stat.washington.edu/thompson/S394/genetics.html>  
[www.youtube.com](http://www.youtube.com) - [www.youtube.com](http://www.youtube.com)  
Википедия - <http://ru.wikipedia.org/wiki/>

## 8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Вероятностные модели генетики" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Мел или маркер, доска

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 01.03.04 "Прикладная математика" и профилю подготовки Математическое моделирование .

Автор(ы):

Володин И.Н. \_\_\_\_\_

Симушкин С.В. \_\_\_\_\_

Салимов Р.Ф. \_\_\_\_\_

"\_\_" \_\_\_\_\_ 201\_\_ г.

Рецензент(ы):

Халиуллин С.Г. \_\_\_\_\_

"\_\_" \_\_\_\_\_ 201\_\_ г.