

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Институт информационных технологий, математики и механики

(факультет / институт / филиал)

УТВЕРЖДЕНО

решением Ученого совета ННГУ

протокол №13 от 30.11.2022 г.

Рабочая программа дисциплины

Математические методы биоинформатики

(наименование дисциплины (модуля))

Уровень высшего образования

бакалавриат

(бакалавриат / магистратура / специалитет)

Направление подготовки / специальность

020302 Фундаментальная информатика и информационные технологии

(указывается код и наименование направления подготовки / специальности)

Направленность образовательной программы

Инженерия программного обеспечения

(указывается профиль / магистерская программа / специализация)

Форма обучения

очная

(очная / очно-заочная / заочная)

Нижегород

2023 год

1. Место дисциплины в структуре ООП

Дисциплина Б1.В.ДВ.07.06 «Математические методы биоинформатики-1» относится к части ООП направления подготовки 020302 Фундаментальная информатика и информационные технологии, формируемой участниками образовательных отношений.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции* (код, содержание индикатора)	Результаты обучения по дисциплине**	
ПК-3 Способен создавать и исследовать новые математические модели в естественных науках, промышленности и бизнесе, с учетом возможностей современных информационных технологий и программирования и компьютерной техники	ПК-3.1: Знает методы анализа и исследования математических моделей в области фундаментальной информатики и информационных технологий;	<i>Знает базовые математические методы исследования для решения прикладных задач биоинформатики</i>	<i>Собеседование</i>
	ПК-3.2: Умеет определять ключевые свойства и ограничения системы	<i>Умеет определять и профессионально реализовывать необходимые для решения прикладных задач биоинформатики математические методы. Умеет применять для исследования пакеты прикладных программ, обрабатывать и анализировать полученные результаты.</i>	<i>Практические задания</i>

3. Структура и содержание дисциплины

3.1. Трудоемкость дисциплины

	Очная форма обучения
Общая трудоемкость	3 ЗЕТ
Часов по учебному плану	108

в том числе	
аудиторные занятия (контактная работа):	25
- занятия лекционного типа	24
- занятия семинарского типа	
- текущий контроль (КСР)	1
самостоятельная работа	83
Промежуточная аттестация –зачет	

3.2. Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)	В том числе				Самостоятельная работа обучающегося, часы
		Контактная работа (работа во взаимодействии с преподавателем), часы. Из них				
		Занятия лекционного типа	Занятия семинарского типа	Занятия лабораторного типа	Всего	
Элементы теории вероятности. – Вероятность. Диаграммы Венна. Условная вероятность и теорема Байеса. Случайные величины и их распределения. Математическое ожидание, моменты и преобразования. Неравенство Маркова. Неравенство Чебышева. Основные дискретные и непрерывные распределения	26	6			6	20
Элементы математической статистики. - Распределения выборки и основные теоремы. Сходимость. Слабый закон больших чисел. Нормальное распределение. Усиленный закон больших чисел. Центральная предельная теорема. Т -распределение, F-распределение. Статистический вывод. Оценка правдоподобия. Байесовская оценка. Приближенные Байесовские вычисления. Марковская цепь Монте-Карло. Гиббсовская выборка. Последовательный Монте-Карло.	26	6			6	20
Элементы статистического анализа. - Проверка статистических гипотез. Z-критерий. Односторонние и двусторонние критерии. Т-критерий. Критерий дисперсии. Ошибки первого и второго рода. Значимость критерия. Проверка гипотез для двух и более выборок. Дисперсионный анализ. Корреляционный и ассоциативный анализ. Критерий χ^2 (критерий Пирсона). Энтропия. Взаимная информация. Линейная корреляция. Взаимная корреляция.	26	6			6	20
Статистический анализ в среде R. – Введение в R. Описательная статистика и графика в R. Перечень тестов и реализация в программной среде R :t-критерий, f-критерий , z-критерий, χ^2 -критерий, критерии Смирнова-Колмогорова, Крускала-Уоллиса, Вилкоксона-Манна-Уинтни, пермутационный критерий для двух выборок. Корреляционный и ассоциативный анализ в среде R: критерий χ^2 (критерий Пирсона). Энтропия. Взаимная информация. Линейная корреляция. Взаимная корреляция.	29	6			6	23
Текущий контроль (КСР)	1				1	

Промежуточная аттестация – зачет						
Итого	108	24			25	83

Текущий контроль успеваемости реализуется в форме опросов на занятиях лекционного типа.

Промежуточная аттестация проходит в традиционной форме (зачет).

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

а. Виды самостоятельной работы студентов

- проработка теоретического материала лекционных занятий;
- подготовка к промежуточной аттестации.

Проработка теоретического материала лекционных занятий

Выполняется самостоятельно обучающимся с использованием материалов лекций и рекомендованной основной и дополнительной литературы. Контроль проводится в часы контроля самостоятельной работы студентов.

Подготовка к практическим занятиям

Выполняется самостоятельно обучающимся с использованием лекционных материалов.

Подготовка к промежуточной аттестации

Фактором успешного прохождения промежуточной аттестации является систематическая работа студента в течение всего периода изучения дисциплины. В этом случае подготовка к зачету будет концентрированной систематизацией всех полученных знаний, умений и навыков.

В качестве методических материалов при подготовке к зачету рекомендуется использовать собственные конспекты лекций и рекомендованную основную и дополнительную литературу.

Выполнение практических заданий на следующие темы:

- 1.1 Введение в R. Описательная статистика и графика в R.
- 1.2 Условная вероятность и теорема Байеса. Случайные величины и их распределения. Математическое ожидание, моменты и преобразования. Неравенство Маркова. Неравенство Чебышева. Основные дискретные и непрерывные распределения.
- 1.3 Распределения выборки и основные теоремы. Сходимость. Слабый закон больших чисел. Нормальное распределение. Усиленный закон больших чисел. Центральная предельная теорема. Т-распределение, F-распределение.
- 1.4 Статистический вывод. Оценка правдоподобия. Байесовская оценка. Приближенные Байесовские вычисления.
- 1.5 Марковская цепь Монте-Карло. Гиббсовская выборка. Последовательный Монте-Карло.
- 1.6 Проверка статистических гипотез. Z-критерий. Односторонние и двусторонние критерии.

Т-критерий.

- 1.7 Критерий дисперсии. Ошибки первого и второго рода. Значимость критерия. Проверка гипотез для двух и более выборок.
- 1.8 Дисперсионный анализ. Перечень тестов и реализация в программной среде R (t-критерий, f-критерий, z-критерий, χ^2 -критерий, критерии Смирнова-Колмогорова, Крускала-Уоллиса, Вилкоксона-Манна-Уитни, пермутационный критерий для двух выборок).
- 1.9 Корреляционный и ассоциативный анализ. Критерий χ^2 (критерий Пирсона).
- 1.10 Энтропия. Взаимная информация. Линейная корреляция. Взаимная корреляция.

b. Образовательные материалы для самостоятельной работы студентов

1. Кремер Н.Ш. Теория вероятностей и математическая статистика:

Учебник для вузов.— М.: ЮНИТИ – ДАНА, 2006.-573 с. (10 экз)

2. Программный пакет R <http://cran.r-project.org/>.

Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине (модулю), включающий:

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности компетенций						
	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	Не зачтено		Зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько незначительных ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минималь-	При решении стандартных	Продемонстрированы	Продемонстрированы все	Продемонстрированы все	Продемонстрированы все	Продемонстрированы все

	ных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	задач не продемонстрированы основные умения. Имели место грубые ошибки.	основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	основные умения, решены все основные задачи с отдельными несущественным недочетами, выполнены все задания в полном объеме.	основные умения, решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами.	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продemonстрирован творческий подход к решению нестандартных задач.

Шкала оценки при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»
	Удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

Критерии оценок за выполнение практического задания (каждое задание оценивается в 1 балл)

Практическое задание выполнено в полном объеме, отчет правильно и аккуратно оформлен	1	Превосходно	Зачтено
		Отлично	
Практическое задание выполнен в полном объеме, но отчет не аккуратно оформлен	0,75	Очень хорошо	
		Хорошо	
Практическое задание выполнено в полном объеме, но не достаточно самостоятельно, отчет оформлен	0,5	Удовлетворительно	
Практическое задание не выполнено	0	Неудовлетворительно	Не зачтено
		Плохо	

5.2. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения

5.2.1 Контрольные вопросы

<i>вопросы</i>	<i>Код формируемой компетенции</i>
1. Операции над множествами. Диаграммы Венна. Законы Де Моргана. Вероятность.	ПК-3
2. Условная вероятность и теорема Байеса. Случайные величины и их распределения.	ПК-3
3. Математическое ожидание, моменты и преобразования. Неравенство Маркова.	ПК-3
4. Неравенство Чебышева. Основные дискретные и непрерывные распределения.	ПК-3
5. Распределения выборки. Сходимость. Слабый закон больших чисел.	ПК-3
6. Усиленный закон больших чисел. Центральная предельная теорема. Т-распределение, F-распределение.	ПК-3
7. Статистический вывод. Оценка правдоподобия.	ПК-3
8. Байесовская оценка. Приближенные Байесовские вычисления.	ПК-3
9. Марковская цепь Монте-Карло. Гиббсовская выборка. Последовательный Монте-Карло.	ПК-3
10. Предмет статистической геномики. Основные задачи.	ПК-3
11. Проверка статистических гипотез. Z-критерий. Односторонние и двусторонние критерии. Т-критерий.	ПК-3
12. Критерий дисперсии. Ошибки первого и второго рода. Значимость критерия. Проверка гипотез для двух и более выборок.	ПК-3
13. Дисперсионный анализ. Основные критерии (t-критерий, f-критерий, z-критерий, χ^2 -критерий, критерии Смирнова-Колмогорова, Крускала-Уоллиса, Вилкоксона-Манна-Уитни, пермутационный критерий для двух выборок).	ПК-3
14. Корреляционный и ассоциативный анализ. Критерий χ^2 (критерий Пирсона).	ПК-3
15. Энтропия. Взаимная информация. Линейная корреляция. Взаимная корреляция.	ПК-3

5.2.2. Типовые практические задания для оценки сформированности компетенции ПК-3

Задание 1. Выполнить базовый статистический анализ набора данных в среде R.

Дано: Набор случайных чисел, распределенных по заданному нормальному закону со средним 20 и дисперсией 1.

Требуется: построить гистограмму и сравнить с теоретическим распределением, вычислить оценку четырех первых моментов.

Задание 2. Оценка правдоподобия в среде R.

Дано: Набор случайных чисел, распределенных по заданному нормальному закону со средним 20 и дисперсией 1.

Требуется: Предположим, что вектор “x” представляет собой случайную выборку.

Протестируйте гипотезу, что среднее равно 20 при условии заданной дисперсии равной 1.

Определите соответствующее р-значение. Предположим, что вектор “x” представляет собой случайную выборку. Протестируйте гипотезу, что среднее равно 20 при условии неизвестной дисперсии, которую вам тоже надо оценить. Определите соответствующее р-значение.

Задание 3. Байесовская оценка в среде R.

Дано: Рассмотрим эксперимент с монеткой, для которой одна сторона “1” выпадает с вероятностью μ , а другая “0” с вероятностью $1 - \mu$. Предположим, что априорное распределение равно 1.

Требуется: Постройте график оценки по Байесу для параметра μ для эксперимента, в котором выпало 001.

Задание 4. Дисперсионный анализ, проверка гипотез в среде R.

Дано: Файл данных с 9 параметрами.

Требуется: Для 9 параметров P1-P9 проанализируйте гипотезу что ряды с пометкой “Abnormal” имеют другое распределение чем ряды с пометкой “Normal”. Используйте тест Kruskal-Wallis. Поместите справа р-значения для тех параметров, для которых нельзя отвергнуть нулевую гипотезу, что распределения одинаковы.

6. Учебно-методическое и информационное обеспечение дисциплины

а) основная литература:

1. Буре, В.М. Теория вероятностей и математическая статистика. [Электронный ресурс] / В.М. Буре, Е.М. Парилина. — Электрон. дан. — СПб. : Лань, 2013. — 416 с. Режим доступа: https://e.lanbook.com/book/10249?category_pk=913#book_name
2. Пугачев, В.С. Теория вероятностей и математическая статистика. [Электронный ресурс] — Электрон. дан. — М. : Физматлит, 2002. — 496 с. — Режим доступа: https://e.lanbook.com/book/48170?category_pk=913#book_name

б) дополнительная литература:

1. Болотюк, В.А. Практикум и индивидуальные задания по курсу теории вероятностей (типовые расчеты). [Электронный ресурс] / В.А. Болотюк, Л.А. Болотюк, А.Г. Гринь, И.П.

Гринь. — Электрон. дан. — СПб.: Лань, 2010. — 288 с. — Режим доступа: https://e.lanbook.com/book/534?category_pk=913#book_name

2. Ибрагимов, И.М. Основы компьютерного моделирования наносистем. [Электронный ресурс] / И.М. Ибрагимов, А.Н. Ковшов, Ю.Ф. Назаров. — Электрон. дан. — СПб.: Лань, 2010. — 384 с. —

Режим доступа: https://e.lanbook.com/book/156?category_pk=915#authors

в) программное обеспечение и Интернет-ресурсы

Программный пакет R <http://cran.r-project.org/> (в свободном доступе), лицензия: <https://www.r-project.org/COPYING>

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения: компьютерный класс, проектор, экран.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду ННГУ.

Лицензионное и свободно распространяемое ПО:

- операционные системы семейства Microsoft Windows, – лицензия по подписке Microsoft Imagine;
- программный пакет R, – <http://cran.r-project.org/> (в свободном доступе), лицензия: <https://www.r-project.org/COPYING>.

Программа составлена в соответствии с требованиями ОС ННГУ 020302 Фундаментальная информатика и информационные технологии

Автор:

проф. каф. прикладной

математики, к.ф.-м.н. _____ Заикин А.А.

Рецензент: _____

Зав. каф. прикладной математики _____ Иванченко М.В.

Программа одобрена на заседании методической комиссии института информационных технологий, математики и механики

от 1 декабря 2021 года, протокол № 2.