

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт математики и механики им. Н.И. Лобачевского



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ

_____ Турилова Е.А.

"__" _____ 20__ г.

Программа дисциплины

Программирование на R

Направление подготовки: 02.03.01 - Математика и компьютерные науки

Профиль подготовки: Наука о данных

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2022

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
 - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
 - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и): доцент, к.н. (доцент) Григорьева И.С. (кафедра математической статистики, Институт математики и механики им. Н.И. Лобачевского), Irina.Grigorieva@kpfu.ru

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-4	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-5	Способен использовать современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

основные объекты, используемые в языке R и действия с ними;
 способы ввода и вывода данных из/в файлы различных форматов;
 набор задач, которые можно решать с помощью языка R;
 способы обновления и расширения возможностей свободного ПО "R".

Должен уметь:

применять векторный вариант объектно-ориентированного программирования;
 решать с помощью R основные задачи статистической обработки данных;
 подключать библиотеки программ для решения широкого класса статистических за-дач.

Должен владеть:

постановки задач статистической обработки;
 поиска в интернет обновлений и расширений языка, подключения их к работе.

Должен демонстрировать способность и готовность:

использовать статистический язык R для решения статистических задач и задач обработки данных
 совершенствовать навыки владения языком с использованием справочных и иных официальных материалов

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.06 Дисциплины (модули)" основной профессиональной образовательной программы 02.03.01 "Математика и компьютерные науки (Наука о данных)" и относится к части ОПОП ВО, формируемой участниками образовательных отношений.

Осваивается на 3 курсе в 6 семестре.

3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

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

Контактная работа - 33 часа(ов), в том числе лекции - 16 часа(ов), практические занятия - 16 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 1 часа(ов).

Самостоятельная работа - 39 часа(ов).

Контроль (зачёт / экзамен) - 0 часа(ов).

Форма промежуточного контроля дисциплины: зачет в 6 семестре.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Се- местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само- стоя- тель- ная ра- бота
			Лекции, всего	Лекции в эл. форме	Практи- ческие занятия, всего	Практи- ческие в эл. форме	Лабора- торные работы, всего	Лабора- торные в эл. форме	
1.	Тема 1. Общее представление о задачах и особенностях среды R. Типы данных языка R. Объекты. Ввод данных с клавиатуры, с помощью таблицы, из файла. Организация обращения к элементам данных	6	4	0	4	0	0	0	7
2.	Тема 2. Вычисление основных характеристик выборки. Вычисление результирующих величин для таблицы (по строкам и столбцам). Группа функций apply()	6	2	0	2	0	0	0	8
3.	Тема 3. Элементы языка программирования. Пользовательские функции Подключение внешних библиотек. Визуализация данных. Использование метода главных компонент. Трехмерная графика.	6	4	0	4	0	0	0	8
4.	Тема 4. Продвинутое методы: линейные модели, регрессионный анализ Кластеризация. Визуализация результатов кластеризации	6	4	0	4	0	0	0	8
5.	Тема 5. Решение контрольных задач	6	2	0	2	0	0	0	8
	Итого		16	0	16	0	0	0	39

4.2 Содержание дисциплины (модуля)

Тема 1. Общее представление о задачах и особенностях среды R. Типы данных языка R. Объекты. Ввод данных с клавиатуры, с помощью таблицы, из файла. Организация обращения к элементам данных

Ввод данных языка R с клавиатуры и из файлов. Результат ввода таких данных в языке приобретает вид таблицы (data.frame). Обращение к элементам векторов и таблиц осуществляется с помощью квадратных скобок [], [[]], по именам полей. Аргументами этих операторов могут быть как числовые, так и логические векторы. Это позволяет отбирать данные по свойствам, в том числе свойствам элементов других векторов/таблиц.

Тема 2. Вычисление основных характеристик выборки. Вычисление результирующих величин для таблицы (по строкам и столбцам). Группа функций apply()

Статистические применения языка R. Первичная обработка набора данных (вектора). Команды summary(), var(), mean(), sd(), median() и т.п.

Специфика векторной безиндексной записи.

Обработка наборов векторов (матриц, таблиц). Группа функций apply (apply, tapply, sapply, lapply).

Создание выборок из различных распределений с последующим подсчетом параметров. Построение доверительных интервалов и сравнение их с истинными значениями.

Тема 3. Элементы языка программирования. Пользовательские функции Подключение внешних библиотек. Визуализация данных. Использование метода главных компонент. Трехмерная графика.

Операторы цикла for, while.

Условные операторы (if, if..else, ifelse и другие)

Пользовательская функция как объект R

Использования пользовательской функции как процедуры и как функции

Использование имен команд как формальных и фактических параметров

Подключение и использование внешних библиотек

Тема 4. Продвинутое методы: линейные модели, регрессионный анализ Кластеризация. Визуализация результатов кластеризации

Команда `lm()` построения линейной модели.

Объект типа `lm()`. Его использование в регрессионном и дискриминантом анализе

Использование команды `summary()` для оценки качества линейной модели

Обобщение: многомерные модели, нелинейные модели

Формулы типа "отклик~воздействие". Использование команды `I()`

Тема 5. Решение контрольных задач

Самостоятельное программирование комплексов программ для полного цикла обработки данных: ввода, обработки и вывода результатов на экран и в файлы.

Примерные темы:

Поиск закономерностей в данных.

Кластеризация многомерных данных различными методами.

Исследование временных рядов и построение прогнозов на их основе.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства науки и высшего образования Российской Федерации от 6 апреля 2021 года №245)

Письмо Министерства образования Российской Федерации №14-55-99бин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

Полезные команды языка R - <https://aakinshin.net/ru/posts/r-functions/>

Сайт языка R - <http://www.r-project.org/>

6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемому результату обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;
- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы. Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;

- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Д.Мертц, Б.Хантинг. Статистическое программирование на R. Часть 1. Купаемся в изобилии статистических возможностей. - <http://www.ibm.com/developerworks/ru/library/l-r1/index.html>

Д.Мертц, Б.Хантинг. Статистическое программирование на R. Часть 2. Функциональное программирование и анализ данных. - <http://www.ibm.com/developerworks/ru/library/l-r3/index.html>.

Д.Мертц, Б.Хантинг. Статистическое программирование на R. Часть 3. Повторное использование кода и объектное программирование. - <http://www.ibm.com/developerworks/ru/library/l-r3/index.html>.

Зеркало для скачивания обновлений - CRAN <http://cran.r-project.org/>

Команда в MS Teams -

<https://teams.microsoft.com/l/team/19%3afacc833b701a4745a987f8f9adac2960%40thread.tacv2/conversations?groupId=0569722f-feb9->

Поиск в материалах по R - R <http://finzi.psych.upenn.edu/nmz.html>

Сайт проекта - <http://www.r-project.org/>

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Вид работ	Методические рекомендации
лекции	Лекции проводятся в аудитории, оснащенной проектором. Преподаватель демонстрирует необходимый материал на экране, попутно проводя опрос среди слушателей на понимание рассказанного материала. Желательно, чтобы студенты также имели доступ к компьютерам, чтобы повторять действия лектора и видеть их результат. Для дистанционного обучения используется команда MS Teams (код команды выдается преподавателем)
практические занятия	Практические занятия проводятся в компьютерном классе, на компьютерах должна быть установлена среда R. Необходим также выход в интернет. В процессе занятия преподаватель напоминает теоретический материал, а также задает студентам задачи для самостоятельного решения. Проводится устный опрос по результатам решения задач.
самостоятельная работа	Установить на домашнем компьютере среду программирования R и иметь копии всех составленных на занятиях программ. Повторить основные понятия математической статистики (оценка параметров, проверка гипотез, дисперсионный анализ, построение линий регрессии и других моделей) Активно пользоваться справочной системой языка. Составить и постоянно пополнять свой список интернет-ресурсов, посвященных языку R. Для консультаций используйте команду MS Teams (код команды выдается преподавателем)

Вид работ	Методические рекомендации
зачет	Зачет проводится в компьютерном классе. Студенту задается вопрос из числа вынесенных на зачет. Он должен не только ответить устно, но и проиллюстрировать свой ответ скриптом, созданным непосредственно во время зачета. Разрешается пользоваться скриптами, составленными студентом в процессе обучения. Преподаватель имеет право задавать дополнительные вопросы, уточняющие ответ студента.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Компьютерный класс.

12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи;
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 02.03.01 "Математика и компьютерные науки" и профилю подготовки "Наука о данных".

Приложение 2
к рабочей программе дисциплины (модуля)
Б1.В.06 Программирование на R

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 02.03.01 - Математика и компьютерные науки

Профиль подготовки: Наука о данных

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2022

Основная литература:

1. Язык и среда программирования R: Учебное пособие / Золотарюк А.В. - М.:НИЦ ИНФРА-М, 2018. - 183 с. - (Высшее образование: Бакалавриат) ISBN 978-5-16-107182-3 (online) - Режим доступа: <http://znanium.com/catalog/product/997099>
2. 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>
3. Буре, В.М. Методы прикладной статистики в R и Excel [Электронный ресурс] : учебное пособие / В.М. Буре, Е.М. Парилина, А.А. Седаков. - Электрон. дан. - Санкт-Петербург : Лань, 2018. - 152 с. - Режим доступа: <https://e.lanbook.com/book/104938>

Дополнительная литература:

1. Шипунов А.Б., Наглядная статистика. Используем R! [Электронный ресурс] / А.Б. Шипунов, Е.М. Балдин, П.А. Волкова, А.И. Коробейников, С.А. Назарова, С.В. Петров, В.Г. Суфиянов. - М. : ДМК Пресс, 2012. - 298 с. - ISBN 978-5-94074-828-1 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785940748281.html>
2. 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>

Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Направление подготовки: 02.03.01 - Математика и компьютерные науки

Профиль подготовки: Наука о данных

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2022

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен обучающимся. ЭБС Издательства "Лань" включает в себя электронные версии книг издательства "Лань" и других ведущих издательств учебной литературы, а также электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. ЭБС Издательства "Лань" обеспечивает доступ к научной, учебной литературе и научным периодическим изданиям по максимальному количеству профильных направлений с соблюдением всех авторских и смежных прав.