

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

_____ Д.А. Таюрский

"__" _____ 20__ г.

Программа дисциплины

Многомерный статистический анализ

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: 2021

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) Симушкин С.В. (кафедра математической статистики, Институт математики и механики им. Н.И. Лобачевского), Sergey.Simushkin@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-2	Способен демонстрировать базовые знания математических и естественных наук, основ программирования и информационных технологий

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

основные вероятностные законы, описывающие поведение случайных векторов (нормальный закон, мультиномиальный закон);

основные принципы построения моделей регрессионных связей между исследуемыми характеристиками (регрессия с линейной зависимостью от неизвестных параметров, полная, частная и множественная корреляция);

основные формы представления многомерных данных и особенности работы с ними;

методы, применяемые для анализа и обработки многомерных совокупностей (полный корреляционный анализ, метод главных компонент, факторный анализ);

Должен уметь:

применять методы многомерного статистического анализа для обработки реальных числовых данных, учитывая границы применимости математической модели;

применять специализированные программные продукты для проведения вычислительных процедур многомерного анализа;

реализовывать процедуры многомерного статистического анализа в рамках имеющихся средств обработки данных;

выбирать методику статистического исследования экспериментальных данных

Должен владеть:

методами построения математической модели профессиональных задач и содержательной интерпретации полученных результатов.

навыками выбора математических методов обработки экспериментальных данных, адекватных целям исследования;

навыками реализации математических методов обработки экспериментальных данных в виде прикладных программных продуктов;

навыками составления отчетов по методикам исследования и их реализации в виде ПО, анализа результатов обработки

Должен демонстрировать способность и готовность:

применять на практике основные методы многомерного статистического анализа

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.05.03 Дисциплины (модули)" основной профессиональной образовательной программы 02.03.01 "Математика и компьютерные науки (Наука о данных)" и относится к дисциплинам по выбору.

Осваивается на 4 курсе в 7 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) единицы(ы) на 72 часа(ов).

Контактная работа - 37 часа(ов), в том числе лекции - 18 часа(ов), практические занятия - 18 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 1 часа(ов).

Самостоятельная работа - 35 часа(ов).

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

Форма промежуточного контроля дисциплины: зачет в 7 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)						Самостоя- тельная 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. Теоретико-вероятностные аспекты многомерного анализа Многомерное нормальное распределение	7	6	0	6	0	0	0	12
2.	Тема 2. Методы сокращения размерности пространства признаков	7	3	0	3	0	0	0	6
3.	Тема 3. Статистические оценки многомерных параметров и их распределения	7	3	0	3	0	0	0	6
4.	Тема 4. Проверка гипотез о параметрах многомерных совокупностей	7	6	0	6	0	0	0	11
	Итого		18	0	18	0	0	0	35

4.2 Содержание дисциплины (модуля)

Тема 1. Теоретико-вероятностные аспекты многомерного анализа Многомерное нормальное распределение

Подготовка данных. Удаление аутлайеров. Вычисление основных числовых характеристик -- вектора средних, матрицы ковариаций, матрицы корреляций. Построение линейного прогноза для каждой характеристики. Графическая иллюстрация -- построение гистограмм, эмпирических функций распределения, эллипсоидов рассеяния. Вычисление частных и канонических корреляций.

Тема 2. Методы сокращения размерности пространства признаков

Определение главных компонент. Построение матрицы нагрузок вектора главных компонент - связь с проблемой собственных решений. Выделение наиболее информативных характеристик. Интерпретация результатов. Потеря информации при сокращении размерности пространства признаков

Построение факторной модели. Проблема вращения. Фундаментальная теорема

Тема 3. Статистические оценки многомерных параметров и их распределения

Оценки по методу моментов - моменты эмпирического распределения. Оценки максимального правдоподобия. Несмещенность и состоятельность оценок. Построение доверительных границ для корреляций, коэффициентов регрессии, поверхности линейной регрессии по реальным данным с использованием средств пакета программ Mathematica.

Тема 4. Проверка гипотез о параметрах многомерных совокупностей

Реализация критерия Фишера для сравнения нескольких групп по реальным данным механической обработки дисков авиатурбин. Анализ точности критерия Фишера методом стохастического моделирования. Распределение статистики Лоули-Хотеллинга, точность хи-квадрат аппроксимации. Распределение статистики лямбда-Уилкса.

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 №14-55-996ин/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

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 планируемым результатам 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)

StatPlus аналитическая статистика - <http://www.statplus.net.ua/ru/>

StatSoft электронный учебник по статистике - <http://www.statsoft.ru/home/textbook/>

Иллюстрированный самоучитель по SPSS - <http://www.learnspss.ru/>

Профессиональный информационно-аналитический ресурс - <http://www.machinelearning.ru/>

Федеральный образовательный портал ЭСМ - <http://ecsocman.hse.ru/>

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

Вид работ	Методические рекомендации
лекции	Лекционный материал содержит большой объем теоретического материала, для освоения которого необходимо производить чтение уже начитанного материала с выделением плохо освоенных частей и решением возникающих задач. Кроме того, необходимо по каждой лекции выделять основные части представленных теорем и утверждений, вокруг которых формировать весь остов теории. Студентам следует стремиться к активизации знаний на занятиях и по другим дисциплинам, которые используются при построении лекции.
практические занятия	Для решения практических задач используются стандартные компьютерные средства обработки статистических данных. Поэтому первый шаг в решении таких задач состоит в доскональном освоении возможностей этих средств обработки. Главное здесь - правильная организация данных, представление их в виде удобной таблицы - фрейма. Интерпретация полученных результатов подразумевает подробное ознакомление с соответствующим теоретическим материалом
самостоятельная работа	Изучение методов многомерного анализа предусматривает систематическую самостоятельную работу студентов над дополнительными материалами; развитие навыков самоконтроля, способствующих интенсификации учебного процесса. Изучение лекционного материала по конспекту лекций должно сопровождаться изучением рекомендуемой литературы, основной и дополнительной. Основной целью организации самостоятельной работы студентов является систематизация и активизация знаний, полученных ими на лекциях и в процессе подготовки к практическим занятиям. Студентам следует стремиться к активизации знаний на занятиях и по другим дисциплинам. Самостоятельная работа по изучению дисциплины предполагает внеаудиторную работу, которая включает решение практических задач различной сложности, рассмотрение вопросов, оставленных на самостоятельное изучение, подготовку к экзамену и зачету.
зачет	При подготовке к зачету обучающемуся рекомендуется повторить весь теоретический материал по соответствующим темам с выявлением ключевых теоретических аспектов и проблем, проработкой дополнительного материала по темам. Особое внимание следует уделить пониманию и интерпретации вводимых определений и их связей с доказываемыми утверждениями. Кроме того необходимо уметь применять приводимые в курсе теоремы и леммы к реальным данным. Лучшему пониманию теоретического материала дисциплины будет способствовать разбор деталей определений, вывода и доказательств утверждений, выявление взаимосвязей между определениями, утверждениями и свойствами объектов, изучаемых в дисциплине. с выявлением возможных следствий и неявных свойств объектов, составлением списка возможных дополнительных вопросов и заданий, подготовку к выполнению практических задач по темам дисциплины.

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.ДВ.05.03 Многомерный статистический анализ

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: 2021

Основная литература:

1. Броневи́ч, А.Г. Анализ неопределенности выделения информативных признаков и представлений изображений [Электронный ресурс] : монография / А.Г. Броневи́ч, А.Н. Каркищенко, А.Е. Лепский. - Электрон. дан. - Москва : Физматлит, 2013. - 320 с. - Режим доступа: <https://e.lanbook.com/book/59666>
2. Королев, В.Ю. Математические основы теории риска [Электронный ресурс] : учебное пособие / В.Ю. Королев, В.Е. Бенинг, С.Я. Шоргин. - Электрон. дан. - Москва : Физматлит, 2011. - 620 с. - Режим доступа: <https://e.lanbook.com/book/2742>
3. Коралов, Л.Б. Теория вероятностей и случайные процессы [Электронный ресурс] / Л.Б. Коралов, Я.Г. Синай ; под ред. Б.М. Гуревича ; пер. с англ. Э.В. Переходцевой. - Электрон. дан. - Москва : МЦНМО, 2014. - 408 с. - Режим доступа: <https://e.lanbook.com/book/71821>
4. Постников, М.М. Аналитическая геометрия [Электронный ресурс] : учебное пособие / М.М. Постников. - Электрон. дан. - Санкт-Петербург : Лань, 2009. - 416 с. - Режим доступа: <https://e.lanbook.com/book/318>

Дополнительная литература:

1. Свешников, А.А. Прикладные методы теории марковских процессов [Электронный ресурс] : монография / А.А. Свешников. - Электрон. дан. - Санкт-Петербург : Лань, 2007. - 192 с. - Режим доступа: <https://e.lanbook.com/book/590>
2. Свешников, А.А. Прикладные методы теории случайных функций [Электронный ресурс] : учебное пособие / А.А. Свешников. - Электрон. дан. - Санкт-Петербург : Лань, 2011. - 464 с. - Режим доступа: <https://e.lanbook.com/book/656>
3. Алон, Н. Вероятностный метод [Электронный ресурс] : учебное пособие / Н. Алон, Д. Спенсер ; под ред. А.А. Сапоженко. - Электрон. дан. - Москва : Издательство 'Лаборатория знаний', 2015. - 323 с. - Режим доступа: <https://e.lanbook.com/book/70705>
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>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.05.03 Многомерный статистический анализ

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: 2021

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.