

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



подписано электронно-цифровой подписью

Программа дисциплины

Дисперсионный анализ в науке о данных

Направление подготовки: 02.04.01 - Математика и компьютерные науки

Профиль подготовки: Статистические методы науки о данных

Квалификация выпускника: магистр

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) 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	Способность проводить научные исследования на основе существующих методов в конкретной области профессиональной деятельности
ПК-3	Способен различным образом представлять и адаптировать математические знания и информационные технологии с учетом уровня аудитории

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

- Теоретические основы дисперсионного анализа
- Основные предпосылки применения методов дисперсионного анализа
- Приёмы обработки реальных данных
- Границы применимости методов дисперсионного анализа

Должен уметь:

- Подготавливать наборы данных к обработке
- Соотносить процедуры дисперсионного анализа с данными
- Интерпретировать результаты статистического анализа

Должен владеть:

- Инструментарием линейного и нелинейного анализа данных
- Приёмами подготовки данных к статистической обработке

Должен демонстрировать способность и готовность:

По окончании курса студент должен уметь

- а) ориентироваться в методах дисперсионного анализа статистических данных;
- б) применять методы дисперсионного анализа к реальным данным;
- в) объяснять полученные результаты.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.03.01 Дисциплины (модули)" основной профессиональной образовательной программы 02.04.01 "Математика и компьютерные науки (Статистические методы науки о данных)" и относится к дисциплинам по выбору части ОПОП ВО, формируемой участниками образовательных отношений.

Осваивается на 2 курсе в 3 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 4 зачетных(ые) единиц(ы) на 144 часа(ов).

Контактная работа - 35 часа(ов), в том числе лекции - 0 часа(ов), практические занятия - 34 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 1 часа(ов).

Самостоятельная работа - 73 часа(ов).

Контроль (зачёт / экзамен) - 36 часа(ов).

Форма промежуточного контроля дисциплины: экзамен в 3 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. Модели дисперсионного анализа и методы планирования экспериментов	3	0	0	8	0	0	0	18
2.	Тема 2. Методы оценивания параметров модели при наличии и отсутствии априорных сведений	3	0	0	8	0	0	0	18
3.	Тема 3. Оптимальное планирование эксперимента	3	0	0	9	0	0	0	19
4.	Тема 4. Непараметрические методы; методы сравнения свободные от распределения	3	0	0	9	0	0	0	18
	Итого		0	0	34	0	0	0	73

4.2 Содержание дисциплины (модуля)

Тема 1. Модели дисперсионного анализа и методы планирования экспериментов

Линейные регрессионные модели и нелинейные регрессионные модели, определеие, свойства, интерпретация. Метод наименьших квадратов, метод максимального правдоподобия. Распределение оценок метода наименьших квадратов. Методы анализа асимптотического распределения оценок. Критерии сравнения планов экспериментов.

Тема 2. Методы оценивания параметров модели при наличии и отсутствии априорных сведений

Линейная и экспоненциальная регрессионные модели. Последовательное планирование экспериментов при наличии и отсутствии априорных сведений. Байесовская парадигма. Пересчет вероятности после наблюдения при наличии априорной информации. Оптимальные планы дисперсионного анализа в байесовском подходе.

Тема 3. Оптимальное планирование эксперимента

Определение понятия информационной матрицы и ее свойства. Линейные критерии оптимальности, их интерпретация. Критерии минимаксного типа, примеры минимаксных критериев. Теорема Кифера-Волфовица. Оптимальные планы проведения эксперимента, критерий оптимальности. Методы поиска глобального и локального экстремумов.

Тема 4. Непараметрические методы; методы сравнения свободные от распределения

Определение понятия ранга и его теоретико-вероятностные свойства и интерпретация. Применение ранговых статистик для задач сравнения групп наблюдений. Преимущества и недостатки ранговых методов в сравнении с классическими подходами. Ранговый дисперсионный анализ со случайными эффектами, его применения.

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

Боровков А.А. Математическая статистика. - СПб.: Лань, 2010. - 704 с. - http://e.lanbook.com/books/element.php?pl1_id=3810

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)

MachineLearning - <http://www.machinelearning.ru/>

StatPlus - http://www.statplus.net.ua/ru/help/source/a_anova_single.htm

StatSoft - электронный учебник по Statistica - <http://www.statsoft.ru/home/textbook/modules/stanman.html>

Анализ экономических данных с использованием Statistica 6,0 - http://cdo.bseu.by/stat1/lab2_1.htm

Класс в Teams -

<https://teams.microsoft.com/l/team/19%3ae3b324edf5524298aaa2945d2cbe8464%40thread.tacv2/conversations?groupId=494ca530-7e1e>

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

Вид работ	Методические рекомендации
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 дисперсионного 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.
экзамен	При подготовке к экзамену обучающемуся рекомендуется повторить весь теоретический материал по соответствующим темам с выявлением ключевых теоретических аспектов и проблем, проработкой дополнительного материала по темам. 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. Важным аспектом по освоению дисциплины является планомерное выполнение всех основных и дополнительных заданий преподавателя. обучающемуся рекомендуется составить план процесса подготовки, включающей изучение, повторение, систематизацию, логическую обработку материала, анализ полученной информации 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 02.04.01 "Математика и компьютерные науки" и магистерской программе "Статистические методы науки о данных".

Приложение 2
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.03.01 Дисперсионный анализ в науке о данных

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 02.04.01 - Математика и компьютерные науки

Профиль подготовки: Статистические методы науки о данных

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2022

Основная литература:

1. Симушкин С. В. Методы теории вероятностей. Ч. I. Многомерные модели. Математические основания: учебное пособие / С. В. Симушкин. - Казань: Изд-во Казан. ун-та, 2016. - 255 с. - URL: https://kpfu.ru/staff_files/F1326184362/Methody_Teor_Ver_I.pdf
2. Буре В. М. и др. Методы прикладной статистики в R и Excel: учебное пособие: 3-е изд., стер. - Санкт-Петербург: Лань, 2022 - 152 с. - URL: <https://e.lanbook.com/book/206042>
3. Герашенко Е. С. и др. Статистическая обработка экспериментальных данных. Дисперсионный и ковариационный анализы в языке R: учебное пособие - Рязань: РГРТУ, 2018 - 32 с. - URL: <https://e.lanbook.com/book/168236>
4. Чураков Е. П. Введение в многомерные статистические методы - Санкт-Петербург: Лань, 2022 - 148 с. - URL: <https://e.lanbook.com/book/212468>
5. 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>
6. Свешников А. А. Прикладные методы теории вероятностей - Санкт-Петербург: Лань, 2022 - 480 с. - URL: <https://e.lanbook.com/book/210821>

Дополнительная литература:

- [1] Коровин Е. Н. и др. Статистические методы обработки биомедицинских данных: учебное пособие для реализации образовательных программ высшего образования бакалавриата по направлению подготовки 'Биотехнические системы и технологии' - Старый Оскол: ТНТ, 2021 - 146, [1] с.
- [2] Ниворожкина Л.И. и др. Многомерные статистические методы в экономике: Учебник: 1 - Москва: Издательский Центр РИОР, 2018 - 203 с. - URL: <http://znanium.com/catalog/document?id=320850>
- [3] Айвазян С.А. Методы эконометрики: Учебник: 1 - Москва: Издательство 'Магистр', 2020 - 512 с. - URL: <http://znanium.com/catalog/document?id=355480>
- [4] Потапова В. Ю. и др. Статистическая обработка экспериментальных данных. Регрессионный анализ в языке R: учебное пособие - Рязань: РГРТУ, 2018 - 52 с. - URL: <https://e.lanbook.com/book/168238>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.03.01 Дисперсионный анализ в науке о данных

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

Направление подготовки: 02.04.01 - Математика и компьютерные науки

Профиль подготовки: Статистические методы науки о данных

Квалификация выпускника: магистр

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 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.