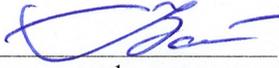


Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования

**«Пермский национальный исследовательский
политехнический университет»**

УТВЕРЖДАЮ

Руководитель программы аспирантуры


Р.А. Файзрахманов
д.э.н., проф., заведующий кафедрой ИТАС

«19» июня 2023 г.

Рабочая программа дисциплины по программе аспирантуры

«Методы и алгоритмы статистического анализа»

Научная специальность	1.2.2. Математическое моделирование, численные методы и комплексы программ
Направленность (профиль) программы аспирантуры	Математическое моделирование, численные методы и комплексы программ
Выпускающая кафедра:	Прикладная физика (ПФ)
Форма обучения	Очная
Курс: 2	Семестр (ы): 4
Виды контроля с указанием семестра:	
Зачет: 4	

Пермь 2023

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
**«Пермский национальный исследовательский
политехнический университет»**

УТВЕРЖДАЮ

Руководитель программы аспирантуры

_____ Р.А. Файзрахманов
д.э.н., проф., заведующий кафедрой ИТАС

«19» июня 2023 г.

Рабочая программа дисциплины по программе аспирантуры

«Методы и алгоритмы статистического анализа»

Научная специальность	1.2.2. Математическое моделирование, численные методы и комплексы программ
Направленность (профиль) программы аспирантуры	Математическое моделирование, численные методы и комплексы программ
Выпускающая кафедра:	Прикладная физика (ПФ)
Форма обучения	Очная
Курс: 2	Семестр (ы): 4
Виды контроля с указанием семестра:	
Зачет: 4	

Пермь 2023

1. Общие положения

Рабочая программа дисциплины «Математическое моделирование, численные методы и комплексы программ» разработана на основании следующих нормативных документов:

- Приказ Минобрнауки России от 20.10.2021 N 951 "Об утверждении федеральных государственных требований к структуре программ подготовки научных и научно-педагогических кадров в аспирантуре (адъюнктуре), условиям их реализации, срокам освоения этих программ с учетом различных форм обучения, образовательных технологий и особенностей отдельных категорий аспирантов (адъюнктов)";
- Постановление Правительства РФ от 30.11.2021 N 2122 "Об утверждении Положения о подготовке научных и научно-педагогических кадров в аспирантуре (адъюнктуре)";
- Самостоятельно устанавливаемые требования к реализуемым программам подготовки научных и научно-педагогических кадров в аспирантуре Пермского национального исследовательского политехнического университета;
- Базовый план по программе аспирантуры;
- Паспорт научной специальности.

1.1 Цель учебной дисциплины

Формирование комплекса знаний, умений и навыков в области математических методов статистической обработки данных с использованием современных компьютерных программ (Microsoft Excel, R, python и др.).

1.2 Место учебной дисциплины в структуре образовательной программы

Дисциплина «Методы и алгоритмы статистического анализа» входит в Блок 1 «Дисциплины» и относится к вариативной части программы. Читается на втором курсе по профилю «Математическое моделирование, численные методы и комплексы программ».

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате изучения дисциплины аспирант должен демонстрировать следующие результаты:

Знать:

- основные понятия и определения теории вероятностей и математической статистики;
- методы анализа данных, часто применяющиеся при статистической обработке результатов;
- признаки «хорошей» оценки.

Уметь:

- правильно поставить задачу эмпирического исследования;
- анализировать полученные результаты, используя методы статистического анализа данных;
- подтвердить или опровергнуть выдвинутые гипотезы;

- применять пакеты программ для решения задач обработки данных и визуализации результатов;
- проводить корреляционный, регрессионный, дисперсионный и кластерный анализ;
- сравнивать параметры распределений нескольких случайных величин, выявлять наличие и вид зависимостей между случайными величинами по данным;
- эффективно применять методы статистического анализа, давать четкую интерпретацию данных.

Владеть:

- статистическими методами обработки данных;
- навыками программирования для разработки программ, которые позволяют автоматизировать процесс анализа данных;
- технологией статистической обработки данных с использованием современных компьютерных программ.

3. Структура учебной дисциплины по видам и формам учебной работы

Таблица 1

Объем и виды учебной работы

№ п.п.	Вид учебной работы	Трудоемкость, ч
		4 семестр
1	Аудиторная работа	21
	В том числе:	
	Лекции (Л)	–
	Практические занятия (ПЗ)	16
2	Контроль самостоятельной работы (КСР)	5
	Самостоятельная работа (СР)	51
	Форма итогового контроля:	Зачет

4. Содержание учебной дисциплины

4.1. Содержание разделов и тем учебной дисциплины

Раздел 1. Часть 1

ПЗ – 6 ч., СРС – 18 ч., КСР – 2 ч.

Тема 1. Основные понятия.

Генеральная совокупность и выборка. Типы переменных. Количественные и номинативные переменные. Меры центральной тенденции. Меры изменчивости. Квартили распределения и график box-plot. Нормальное распределение. Центральная предельная теорема. Доверительные интервалы для среднего. р-уровень значимости.

Тема 2. Сравнение средних.

T-распределение. Сравнение двух средних. t-критерий Стьюдента. Проверка распределения на нормальность. Однофакторный дисперсионный анализ. Множественные сравнения.

Тема 3. Корреляция и регрессия.

Понятие корреляции. Условия применения коэффициента корреляции. Регрессия с одной независимой переменной. Гипотеза о значимости взаимосвязи и коэффициент детерминации. Условия применения линейной регрессии с одним предиктором. Применение регрессионного анализа и интерпретация результатов. Задача предсказания значений зависимой переменной. Регрессионный анализ с несколькими независимыми переменными. Выбор наилучшей модели. Классификация: логистическая регрессия и кластерный анализ.

Раздел 2. Часть 2

ПЗ – 6 ч., СРС – 18 ч., КСР – 2 ч.

Тема 4. Анализ номинативных данных.

Постановка задачи. Расстояние Пирсона. Распределение Хи-квадрат Пирсона. Расчет p-уровня значимости. Анализ таблиц сопряженности. Точный критерий Фишера.

Тема 5. Логистическая регрессия и непараметрические методы.

Логистическая регрессия. Модель без предикторов. Модель с одним номинативным предиктором. Модель с двумя номинативными предикторами. Взаимодействие номинативных предикторов. Когда нужно использовать параметрические методы? Непараметрические методы.

Тема 6. Кластерный анализ и метод главных компонент.

Кластерный анализ методом k-средних. Как определить оптимальное число кластеров? Иерархическая кластеризация. Введение в метод анализа главных компонент.

Раздел 3. Часть 3

ПЗ – 4 ч., СРС – 15 ч., КСР – 1 ч.

Тема 7. Подробнее о линейной регрессии.

Линейность взаимосвязи. Логарифмическая трансформация переменных. Проблема гетероскедастичности. Мультиколлинеарность.

Тема 8. Смешанные регрессионные модели.

Нарушения допущения о независимости наблюдений. Смешанные регрессионные модели (реализация в R или python). Статистическая значимость, обобщенные модели и случайные эффекты. Jackknife и Bootstrap в анализе данных.

4.2. Перечень тем практических занятий

Таблица 2

Темы практических занятий (из пункта 4.1)

№ п.п.	Номер темы дисциплины	Наименование темы практического занятия	Наименование оценочного средства	Представление оценочного средства
1	1	Вычисление среднего, медианы и построение графика box-plot на различных тестовых данных. Нормальное распределение.	Собеседование.	Вопросы по темам / разделам дисциплины.
2	2	Выполнить сравнение двух распределений, используя t-	Собеседование.	Вопросы по темам / разделам

		критерий Стьюдента и провести их проверку на нормальность.		дисциплины.
3	3	Провести корреляционный и регрессионный анализ. Разработать модель регрессионного анализа с несколькими переменными.	Собеседование.	Вопросы по темам / разделам дисциплины.
4	5	Построить модель логистической регрессии.	Собеседование.	Вопросы по темам / разделам дисциплины.
5	6	Разработать модель кластерного анализа методом k-средних (часть 1).	Собеседование.	Вопросы по темам / разделам дисциплины.
6	6	Разработать модель кластерного анализа методом k-средних (часть 2).	Собеседование.	Вопросы по темам / разделам дисциплины.
7	8	Выполнить логарифмическую трансформацию переменных и дать интерпретацию результатам.	Собеседование.	Вопросы по темам / разделам дисциплины.
8	8	Реализовать смешанные регрессионные модели.	Собеседование.	Вопросы по темам / разделам дисциплины.

4.3. Перечень тем для самостоятельной работы аспирантов

Самостоятельная работа аспирантов заключается в теоретическом изучении конкретных вопросов и выполнении творческих заданий.

Таблица 3

Темы самостоятельных заданий

№ п.п.	Номер темы дисциплины	Наименование темы самостоятельной работы	Наименование оценочного средства	Представление оценочного средства
1	3	1. Регрессионный анализ с несколькими независимыми переменными. 2. Выбор наилучшей модели. 3. Классификация: логистическая регрессия и кластерный анализ.	Собеседование	Вопросы по темам / разделам дисциплины
2	5	1. Когда нужно использовать параметрические методы? 2. Непараметрические методы.	Собеседование.	Вопросы по темам / разделам дисциплины.
3	6	1. Кластерный анализ методом k-средних. 2. Как определить оптимальное	Собеседование.	Вопросы по темам / разделам дисциплины.

		число кластеров?		
4	8	1. Статистическая значимость, обобщенные модели и случайные эффекты. 2. Jackknife и Bootstrap в анализе данных.	Собеседование	Вопросы по темам / разделам дисциплины

5. Методические указания для аспирантов по изучению дисциплины

При изучении дисциплины «Методы и алгоритмы статистического анализа» аспирантам целесообразно выполнять следующие рекомендации:

1. изучение учебной дисциплины должно вестись систематически;
2. после изучения какого-либо раздела по учебнику или конспектным материалам рекомендуется по памяти воспроизвести основные термины, определения, понятия раздела;
3. вся тематика вопросов, изучаемых самостоятельно, задается преподавателем. Им же даются источники (в первую очередь вновь изданные в периодической научной литературе) для более детального понимания вопросов, озвученных на лекции.

6. Перечень учебно-методического, библиотечно-справочного и информационного, информационно-справочного обеспечения для работы аспиранта по дисциплине

6.1. Библиотечные фонды и библиотечно-справочные системы

№	Библиографическое описание (автор, заглавие, вид издания, место, издательство, год издания, количество страниц)	Количество экземпляров в библиотеке+кафедре; местонахождение электронных изданий
1	2	3
1 Основная литература		
1	Мхитарян, В.С. Анализ данных: учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2023. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/511020 .	Образовательная платформа «ЮРАЙТ» https://elib.pstu.ru/Record/RUURAIT511020
2	Демидова, Л. А. Разведочный анализ данных. Python : учебно-методическое пособие / Л. А. Демидова. — Москва : РТУ МИРЭА, 2022 — Часть 1 — 2022. — 107 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/310970 .	ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-310970

№	<p align="center">Библиографическое описание (автор, заглавие, вид издания, место, издательство, год издания, количество страниц)</p>	<p>Количество экземпляров в библиотеке+кафедре; местонахождение электронных изданий</p>
1	2	3
3	<p>Волкова, В.М. Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семенова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск : Новосибирский государственный технический университет, 2017. — 74 с. — ISBN 978-5-7782-3183-2. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/91682.html.</p>	<p>ЭБС «ПНИПУ» https://elib.pstu.ru/Record/RUIPRSMART91682</p>
4	<p>Ганичева, А. В. Математическая статистика : учебное пособие / А. В. Ганичева. — Тверь : Тверская ГСХА, 2018. — 176 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/134089.</p>	<p>ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-134089</p>
2 Дополнительная литература		
2.1 Учебно-методические, научные издания		
1	<p>Храмов, А. Г. Методы и алгоритмы интеллектуального анализа данных : учебное пособие / А. Г. Храмов. — Самара : Самарский университет, 2019. — 176 с. — ISBN 978-5-7883-1414-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/148603.</p>	<p>ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-148603</p>
2	<p>Дронов, С. В. Методы и задачи многомерной статистики : монография / С. В. Дронов. — Барнаул : АлтГУ, 2015. — 275 с. — ISBN 978-5-7904-1957-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/154930.</p>	<p>ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-154930</p>
3	<p>Абилова, Е. В. Статистические методы обработки данных на основе информационных технологий : учебное пособие / составитель Е. В. Абилова. — Челябинск : ЮУТУ, 2020. — 106 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/177111.</p>	<p>ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-177111</p>
4	<p>Тарасов, И. Е. Статистический анализ данных в информационных системах : учебно-методическое пособие / И. Е. Тарасов. — Москва : РТУ МИРЭА, 2020. — 96 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/163854.</p>	<p>ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-163854</p>

№	Библиографическое описание (автор, заглавие, вид издания, место, издательство, год издания, количество страниц)	Количество экземпляров в библиотеке+кафедре; местонахождение электронных изданий
1	2	3
5	Карабутов, Н. Н. Введение в теорию эксперимента в исследовании систем / Н. Н. Карабутов. — Санкт-Петербург : Лань, 2023. — 168 с. — ISBN 978-5-507-44900-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/276650 .	ЭБС «Лань» https://elib.pstu.ru/Record/RULANRU-LAN-BOOK-276650
2.2 Периодические издания		
1	Не требуются.	
2.3 Нормативно-технические издания		
1	Не требуются.	
2.4 Официальные издания		
1	Не требуются.	

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

6.2.1. Информационные и информационно-справочные системы

1. Электронная библиотека Пермского национального исследовательского политехнического университета [Электронный ресурс] : [полнотекстовая база данных электрон. док., издан. в Изд-ве ПНИПУ] / Перм. нац. исслед. политехн. ун-т, Науч. б-ка. — Пермь, 2016. — Режим доступа: <http://elib.pstu.ru>, свободный. — Загл. с экрана.

2. Электронно-библиотечная система Издательство «Лань» [Электронный ресурс] : [полнотекстовая база данных: электрон. версии кн., журн. по гуманитар., обществ., естеств. и техн. наукам] / Электрон.-библ. система «Изд-ва «Лань». — Санкт-Петербург, 2010-2016. — Режим доступа : <http://e.lanbook.com>, по IP-адресам компьютер. сети Перм. нац. исслед. политехн. ун-та. — Загл. с экрана.

3. ProQuest Dissertations & Theses Global [Электронный ресурс] : [полнотекстовая база данных : дис. и дипломные работы на ин. яз. по всем отраслям знания] / ProQuest LLC. — Ann Arbor, 2016. — Режим доступа : <http://search.proquest.com/pqdtglobal/dissertations>, по IP-адресам компьютер. сети Перм. нац. исслед. политехн. ун-та. — Загл. с экрана.

4. Электронная библиотека диссертаций РГБ [Электронный ресурс] : [полнотекстовая база данных: электрон. версии дис. и автореф. дис. по всем отраслям знания] / Электрон. б-ка дис. — Москва, 2003-2016. — Режим доступа : <http://diss.rsl.ru>, компьютер. сеть Науч. б-ки Перм. нац. исслед. политехн. ун-та. — Загл. с экрана.

5. Cambridge Journals [Electronic resource : полнотекстовая база данных : электрон. журн. по гуманитар., естеств., и техн. наукам на англ. яз.] / University of Cambridge. — Cambridge : Cambridge University Press, 1770-2012. — Режим доступа : <http://journals.cambridge.org/>. — Загл. с экрана.

6. Справочная Правовая Система КонсультантПлюс [Электронный ресурс] : [полнотекстовая база данных правовой информ. : док., коммент., кн., ст., обзоры и др.]. —

7. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

7.1. Основное учебное оборудование. Рабочее место аспиранта.

Таблица 4

№ п.п.	Наименование и марка оборудования (стенда, макета, плаката, лабораторное оборудование)	Кол-во ед.	Форма приобретения / владения (собственность, оперативное управление, аренда и т.п.)	Номер аудитории
1	2	3	4	5
1	Персональные компьютеры (локальная компьютерная сеть)	5	Оперативное управление	124

8. Фонд оценочных средств

Освоение учебного материала дисциплины запланировано в течение одного семестра. Формой контроля освоения результатов обучения по дисциплине является зачет, проводимый с учетом результатов текущего контроля.

8.1. Описание показателей и критериев оценивания, описание шкал оценивания.

Контроль качества освоения дисциплины включает в себя текущий контроль успеваемости и промежуточную аттестацию аспирантов.

Текущий контроль

Текущий контроль успеваемости обеспечивает оценку освоения дисциплин и проводится в форме собеседования.

- **Собеседование**

Для оценки **знаний** аспирантов проводится собеседование в виде специальной беседы преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной для выяснения объема знаний по определенному разделу, теме, проблеме.

Собеседование может выполняться в индивидуальном порядке или группой аспирантов.

Промежуточная аттестация

Допуск к промежуточной аттестации осуществляется по результатам текущего контроля. Промежуточная аттестация проводится в виде зачета по дисциплине.

- **Шкалы оценивания результатов обучения при сдаче зачета:**

Оценка результатов обучения по дисциплине проводится по шкале оценивания «зачтено», «незачтено» путем выборочного контроля во время зачета.

Шкалы и критерии оценки результатов обучения при сдаче зачета приведены в табл. 5.

Таблица 5

Шкала оценивания результатов на зачете

Оценка	Критерии оценивания
<i>Зачтено</i>	<p>Аспирант продемонстрировал сформированные и систематические знания при ответе на теоретический вопрос билета. Показал отличные знания в рамках усвоенного учебного материала. Ответил на все или большинство дополнительных вопросов.</p> <p>Аспирант правильно выполнил контрольное задание билета. Показал успешное и систематическое применение полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала. Ответил на все или большинство дополнительных вопросов.</p>
<i>Незачтено</i>	<p>При ответе на теоретический вопрос билета аспирант продемонстрировал фрагментарные знания. При ответах на дополнительные вопросы было допущено множество неправильных ответов.</p> <p>При выполнении контрольного задания билета аспирант продемонстрировал частично освоенное умение и применение полученных навыков при решении профессиональных задач в рамках усвоенного учебного материала. При ответах на дополнительные вопросы было допущено множество неточностей.</p>

9. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине

Задания для текущего контроля и проведения промежуточной аттестации должны быть направлены на оценивание:

1. уровня освоения теоретических понятий, научных основ профессиональной деятельности;
2. степени готовности аспиранта применять теоретические знания и профессионально значимую информацию, и оценивание сформированности когнитивных умений.
3. приобретенных умений, профессионально значимых для профессиональной деятельности.

10. Типовые контрольные вопросы и задания или иные материалы, необходимые для оценки результатов освоения дисциплины

Перечень контрольных вопросов и заданий для сдачи зачета по дисциплине «Методы и алгоритмы статистического анализа», преподаваемой в рамках программы аспирантуры по научной специальности 1.2.2 «Математическое моделирование, численные методы и комплексы программ», разработан с учетом научных достижений научно-исследовательской деятельности кафедры.

Типовые творческие задания:

1. Приведите пример реализации регрессионного анализа с использованием современных компьютерных программ.
2. Проведите расчет p -уровня значимости.
3. Выполните анализ данных, используя таблицы сопряженности.

Типовые контрольные вопросы для оценивания знаний на зачете по дисциплине:

1. Что такое меры центральной тенденции и для чего они используются?
2. Квартили распределения (зачем нужны, как их рассчитать)?
3. Кластерный анализ методом k -средних.

Типовые контрольные задания для оценивания приобретенных умений на зачете по дисциплине:

1. По выданному набору данных реализовать box-plot в программной среде R.
2. Перечислить отличия параметрических методов, выделить их преимущества и недостатки.
3. Опишите методику определения оптимального числа кластеров при проведении кластерного анализа методом k -средних.

Полный комплект вопросов и заданий в форме утвержденных билетов хранится на кафедре «ПФ».

МИНОБРНАУКИ РОССИИ
ФГАОУ ВО «Пермский национальный
исследовательский политехнический
университет» (ПНИПУ)

Программа
Математическое моделирование,
численные методы и комплексы программ

Кафедра
Прикладная физика

Дисциплина
«Методы и алгоритмы статистического анализа»

БИЛЕТ № 1

1. Сравнение средних.
2. Логистическая регрессия и непараметрические методы.
3. Обозначить этапы разработки модели кластерного анализа методом k-средних.

Составитель _____
(подпись)

Красняков И.В.

Заведующий кафедрой _____
(подпись)

Брацун Д.А.

«__» _____ 20__ г.

Лист регистрации изменений

№ п.п.	Содержание изменения	Дата, номер протокола заседания кафедры. Подпись заведующего кафедрой
1	2	3
1		
2		
3		
4		