АВТОНОМНАЯ НЕКОММЕРЧЕСКАЯ ОБРАЗОВАТЕЛЬНАЯ ОРГАНИЗАЦИЯ ВЫСШЕГО ОБРАЗОВАНИЯ «НАУЧНО-ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ «СИРИУС» (АНОО ВО «УНИВЕРСИТЕТ «СИРИУС»)

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 0EE1DD57C56F53844743BD8E019F41EF
Владелец: Некрасова Людмила Андреевна
Действителен: с 08.07.2026 до 08.07.2027

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Статистическая обработка данных научных исследований»

Шифр и наименование области науки:	1. Естественные науки
Шифр и наименование группы научных специальностей:	1.2. Компьютерные науки и информатика
Шифр и наименование научной специальности:	1.2.2. Математическое моделирование, численные методы и комплексы программ
Форма обучения:	Очная
Срок освоения образовательной программы:	3 года
Год начала освоения образовательной программы:	2026
Структурное подразделение, ответственное за реализацию образовательной программы:	Научный центр информационных технологий и искусственного интеллекта

АННОТАЦИЯ
к рабочей программе дисциплины (модулю)

Статистическая обработка данных научных исследований, трудоемкость 3 з.е.

1.1. Цель освоения и краткое содержание дисциплины (модуля)

Цель:

Дисциплина «Статистическая обработка данных научных исследований» направлена на формирование у аспирантов профессиональных компетенций в области статистической обработки экспериментальных данных научных исследований с использованием компьютерных технологий.

Краткое содержание (тематика):

Тема 1. Основы статистической обработки данных
Понятие выборки и генеральной совокупности.
Типы данных и шкал измерения.
Программное обеспечение для статистического анализа (обзор).
Тема 2. Описательная статистика
Меры центральной тенденции: среднее, медиана, мода.
Меры разброса: дисперсия, стандартное отклонение, квартили.
Графическое представление данных: гистограммы, box-plots, диаграммы рассеяния.
Тема 3. Проверка статистических гипотез
Нулевая и альтернативная гипотезы.
Уровни значимости и p-значения.
Параметрические и непараметрические критерии (t-критерий Стьюдента, U-критерий Манна-Уитни, $\chi 2$-критерий Пирсона).
Тема 4. Корреляционный и регрессионный анализ
Коэффициент корреляции Пирсона и Спирмена.
Линейная регрессия: построение модели и оценка её качества.
Множественная регрессия.
Тема 5. Дисперсионный анализ
Однофакторный и многофакторный дисперсионный анализ (ANOVA).
Post-hoc тесты (критерий Тьюки).
Тема 6. Многомерные методы анализа данных
Факторный анализ.
Кластерный анализ.
Метод главных компонент (PCA).
Тема 7. Визуализация и представление результатов
Правила оформления таблиц и графиков для научных публикаций.
Подготовка презентаций и докладов.

1.2. Планируемые результаты обучения по дисциплине (модулю)

1.2.1. Сдан зачет по дисциплине «Статистическая обработка данных научных исследований»;

1.2.2. Сформированы личностные компетенции, позволяющие:

- выбирать подходящие методы статистического анализа для конкретных задач исследования;
- обрабатывать экспериментальные данные с помощью специализированного ПО;
- интерпретировать результаты статистического анализа;
- представлять результаты исследования в виде таблиц, графиков и диаграмм.

1. Общая характеристика дисциплины

1.1. Цель дисциплины:

Дисциплина «Статистическая обработка данных научных исследований» направлена на формирование у аспирантов профессиональных компетенций в области статистической обработки экспериментальных данных научных исследований с использованием компьютерных технологий.

1.2. Задачи дисциплины:

- изучить основные понятия и термины статистических методов;
- освоить этапы статистической обработки данных;
- научиться применять методы математической статистики для решения задач обработки данных;
- познакомиться со специализированным программным обеспечением для статистического анализа и визуализации данных;
- сформировать знания о моделях и способах представления экспериментальных данных.

1.3. Место дисциплины в структуре программы аспирантуры:

Дисциплина «Статистическая обработка данных научных исследований» входит в образовательный компонент образовательных программ высшего образования – программ подготовки научных и научно-педагогических кадров в аспирантуре.

Дисциплина является вариативной.

Освоение курса «Статистическая обработка данных научных исследований» опирается на знания, умения, навыки и компетенции, сформированные на двух предшествующих уровнях образования.

1.4. Общая трудоемкость дисциплины: 3 з.е.

1.5. Планируемые результаты обучения по дисциплине:

1.5.1. Сдан зачет по дисциплине «Статистическая обработка данных научных исследований»;

1.5.2. Сформированы личностные компетенции, позволяющие:

- выбирать подходящие методы статистического анализа для конкретных задач исследования;
- обрабатывать экспериментальные данные с помощью специализированного ПО;
- интерпретировать результаты статистического анализа;
- представлять результаты исследования в виде таблиц, графиков и диаграмм.

2. Структура и содержание дисциплины

2.1. Объем дисциплины и виды учебной деятельности:

Виды учебной деятельности	Всего
Контактная работа обучающихся с преподавателем, ак. ч.	22
Самостоятельная работа обучающихся, ак. ч.	86
Промежуточная аттестация, ак. ч.	2
Общая трудоемкость, ак. ч.	108
Общая трудоемкость, з.е.	3

2.2. Структура дисциплины по разделам (темам) и видам учебной деятельности:

Наименования разделов (тем) дисциплины	Контактная работа, ак. ч.	Самостоятельная работа, ак. ч.	Всего, ак. ч.	Форма текущего контроля / промежуточной аттестации
Тема 1. Основы статистической обработки данных	2	12	14	Устные опросы
Тема 2. Описательная статистика	2	12	14	Устные опросы Практические работы
Тема 3. Проверка статистических гипотез	2	12	14	Устные опросы Контрольная работа
Тема 4. Корреляционный и регрессионный анализ	4	12	14	Устные опросы Отчёт по исследованию
Тема 5. Дисперсионный анализ	2	12	14	Устные опросы
Тема 6. Многомерные методы анализа данных	4	12	16	Устные опросы Отчёт по исследованию
Тема 7. Визуализация и представление результатов	4	14	18	Устные опросы Презентация результатов
Промежуточная аттестация	2	-	2	Зачет
Итого	22	86	108	

2.3. Содержание разделов (тем) дисциплины:

Наименования разделов (тем) дисциплины	Содержание разделов (тем) дисциплины
Тема 1. Основы статистической обработки данных	Понятие выборки и генеральной совокупности. Типы данных и шкал измерения. Программное обеспечение для статистического анализа (обзор).
Тема 2. Описательная статистика	Меры центральной тенденции: среднее, медиана, мода. Меры разброса: дисперсия, стандартное отклонение, квартили. Графическое представление данных: гистограммы, box-plots, диаграммы рассеяния.
Тема 3. Проверка статистических гипотез	Коэффициент корреляции Пирсона и Спирмена. Линейная регрессия: построение модели и оценка её качества. Множественная регрессия.
Тема 4. Корреляционный и регрессионный анализ	Коэффициент корреляции Пирсона и Спирмена. Линейная регрессия: построение модели и оценка её качества. Множественная регрессия.

Тема 5. Дисперсионный анализ	Однофакторный и многофакторный дисперсионный анализ (ANOVA). Post-hoc тесты (критерий Тьюки).
Тема 6. Многомерные методы анализа данных	Факторный анализ. Кластерный анализ. Метод главных компонент (PCA).
Тема 7. Визуализация и представление результатов	Правила оформления таблиц и графиков для научных публикаций. Подготовка презентаций и докладов.

3. Текущий контроль и промежуточная аттестация по дисциплине.
Оценочные материалы

3.1. Текущий контроль успеваемости по дисциплине «Управление научными и научно – технологическими проектами» проводится в течение семестра в следующих формах:

Наименования разделов (тем) дисциплины	Форма текущего контроля	Оценочные материалы
Тема 1. Основы статистической обработки данных	Устные опросы	Вопросы для устных опросов
Тема 2. Описательная статистика	Устные опросы Практические работы	Вопросы для устных опросов Задания для практических работ
Тема 3. Проверка статистических гипотез	Устные опросы Контрольная работа	Вопросы для устных опросов Задагия для контрольной работы
Тема 4. Корреляционный и регрессионный анализ	Устные опросы Отчёт по исследованию	Вопросы для устных опросов Задание для подготовки отчёта по исследованию
Тема 5. Дисперсионный анализ	Устные опросы	Вопросы для устных опросов
Тема 6. Многомерные методы анализа данных	Устные опросы Отчёт по исследованию	Вопросы для устных опросов Задание для подготовки отчёта по исследованию
Тема 7. Визуализация и представление результатов	Устные опросы Презентация результатов	Вопросы для устных опросов Задания для подготовки презентации результатов

3.2. Оценочные материалы для текущего контроля:

1. Устные опросы (по темам 1–7)

Примеры вопросов:

Тема 1. Основы статистической обработки данных:

Что такое генеральная совокупность и выборка? В чём их различие?

Какие типы данных существуют в статистике? Приведите примеры.

Назовите 2–3 программных пакета для статистического анализа и кратко охарактеризуйте их возможности.

Тема 2. Описательная статистика:

Как рассчитать среднее арифметическое, медиану и моду? В каких случаях предпочтительнее использовать каждый показатель?
Что характеризуют дисперсия и стандартное отклонение?
Какие виды графиков подходят для визуализации количественных и категориальных данных?
Тема 3. Проверка статистических гипотез:
Сформулируйте нулевую и альтернативную гипотезы для сравнения средних двух групп.
Что означает p-значение? При каком p-значении гипотеза обычно отвергается?
В чём разница между параметрическими и непараметрическими критериями? Приведите по одному примеру.

2. Практические работы

Задание для практической работы по теме 2 (описательная статистика):
Даны результаты измерения роста (см) у 20 студентов: 170,165,180,172,168,175,173,169,171,174,176,167,178,170,172,177,166,179,173,175.
Требуется:
Рассчитать среднее арифметическое, медиану, моду.
Найти дисперсию и стандартное отклонение.
Построить гистограмму распределения роста в программе SPSS/R/Python.
Сделать вывод о распределении признака (симметричное/асимметричное, наличие выбросов).
Критерии оценки:
правильность расчётов — до 5 баллов;
корректность построения графика — до 3 баллов;
обоснованность выводов — до 2 баллов.

3. Контрольная работа (по теме 3)

Вариант 1:
Сформулируйте H0 и H1 для проверки гипотезы о различии средних значений артериального давления у пациентов до и после лечения.
Выберите подходящий статистический критерий для проверки этой гипотезы и обоснуйте выбор.
Даны p-значения: 0,03, 0,15, 0,001. Для каждого укажите, отвергается ли H0 при уровне значимости $\alpha=0,05$.
Выполните расчёт t-критерия Стьюдента для двух независимых выборок (данные предоставляются) в программе R/SPSS.
Критерии оценки:
за каждый правильный ответ на теоретические вопросы — до 2 баллов;
за корректный расчёт и интерпретацию — до 4 баллов.

4. Отчёт по исследованию (темы 4 и 6)

Задание:
Проанализируйте зависимость успеваемости студентов (Y, средний балл) от количества часов самостоятельной работы в неделю (X).
Данные:

Студент	Часы (X)	Балл (Y)
1	5	3,8
2	8	4,2

Студент	Часы (X)	Балл (Y)
…	…	…
30	12	4,7

Требования к отчёту:
Постройте диаграмму рассеяния X и Y.
Рассчитайте коэффициент корреляции Пирсона.
Постройте линейную регрессионную модель Y=a+bX.
Оцените качество модели (коэффициент детерминации R2).
Сделайте выводы о силе и направлении связи, дайте прогноз среднего балла при X=10 часов.
Критерии оценки:
визуализация — до 2 баллов;
корректность расчётов — до 4 баллов;
интерпретация результатов — до 4 баллов.

5. Презентация результатов (тема 7)

Задание:
Подготовьте слайд-презентацию (5–7 слайдов) по результатам дисперсионного анализа (ANOVA) для сравнения эффективности трёх методов обучения.
Содержание слайдов:
Цель и гипотеза исследования.
Описание данных и методов анализа.
Таблица с результатами ANOVA (F-статистика, p-значение).
График средних значений с доверительными интервалами.
Выводы и рекомендации.
Критерии оценки:
структура презентации — до 2 баллов;
наглядность графиков и таблиц — до 3 баллов;
логичность выводов — до 3 баллов;
качество оформления — до 2 баллов.

Шкала оценивания текущего контроля

«Отлично»: полное выполнение задания, глубокие знания теории, корректные расчёты, чёткие выводы.
«Хорошо»: незначительные ошибки в расчётах или интерпретации, неполнота выводов.
«Удовлетворительно»: существенные ошибки в расчётах, слабая аргументация выводов.
«Неудовлетворительно»: задание не выполнено или выполнено с грубыми ошибками.

Вопросы для зачёта по дисциплине «Статистическая обработка данных научных исследований»

Раздел 1. Основы статистики и описательный анализ

Что такое генеральная совокупность и выборка? В чём их принципиальное различие?
Перечислите типы данных в статистике (номинальные, порядковые, интервальные, отношения) и приведите примеры для каждого типа.
Какие существуют шкалы измерения данных? Кратко охарактеризуйте каждую.

Назовите и опишите основные меры центральной тенденции (среднее, медиана, мода). В каких случаях предпочтительно использовать каждую из них?
Какие показатели характеризуют разброс данных? Дайте определение дисперсии, стандартного отклонения и коэффициента вариации.
Какие графические методы используются для визуализации количественных и категориальных данных? Приведите 2–3 примера для каждого случая.
Раздел 2. Проверка статистических гипотез
Что такое нулевая (H0) и альтернативная (H1) гипотезы? Приведите пример формулировки для сравнения средних двух групп.
Что означает уровень значимости α? Какое стандартное значение α используется в научных исследованиях?
Что такое p-значение? Как оно используется при проверке гипотез?
В чём разница между параметрическими и непараметрическими критериями? Приведите по одному примеру каждого типа.
Опишите условия применения t-критерия Стьюдента. Какие существуют разновидности этого критерия?
Для каких задач используется χ2-критерий Пирсона? Каковы его ограничения?
Как интерпретировать результаты проверки гипотез: когда H0 отвергается, а когда нет?
Раздел 3. Корреляционный анализ
Что показывает коэффициент корреляции? Какие значения он может принимать?
В чём различие между коэффициентом корреляции Пирсона и Спирмена? Когда какой использовать?
Как интерпретировать силу и направление корреляционной связи? Приведите шкалу интерпретации (например, по Эвансу).
Может ли корреляция указывать на причинно-следственную связь? Обоснуйте ответ.
Раздел 4. Регрессионный анализ
В чём состоит суть регрессионного анализа? Чем он отличается от корреляционного?
Опишите линейную регрессионную модель: $Y=a+bX$. Что означают параметры a и b?
Что такое метод наименьших квадратов (МНК)? Как он используется в регрессионном анализе?
Что показывает коэффициент детерминации $R2$? Как его интерпретировать?
Какие предположения должны выполняться для корректного применения линейной регрессии?
Что такое множественная регрессия? Какие проблемы могут возникнуть при её использовании (мультиколлинеарность и др.)?
Раздел 5. Дисперсионный анализ (ANOVA)
Для каких задач применяется дисперсионный анализ? В чём его основное преимущество перед t-тестами?
Опишите однофакторный дисперсионный анализ. Какие гипотезы проверяются в этом случае?
Что такое F-статистика в ANOVA? Как её интерпретировать?
Зачем нужны post-hoc тесты (например, критерий Тьюки)? Приведите пример ситуации, когда они необходимы.
В чём отличие однофакторного и многофакторного ANOVA?
Раздел 6. Многомерные методы анализа
Для чего используется факторный анализ? Кратко опишите его суть.
Что такое метод главных компонент (PCA)? Как он помогает в анализе данных?

В чём состоит задача кластерного анализа? Какие методы кластеризации вы знаете?
Как оценить качество кластеризации? Приведите 1–2 показателя.
Раздел 7. Практические аспекты и программное обеспечение
Назовите 3–4 программных пакета для статистического анализа данных. Кратко сравните их возможности.
Какие правила оформления таблиц и графиков следует соблюдать в научных публикациях?
Как правильно представить результаты статистического анализа в научном отчёте или статье? Перечислите основные элементы.
Какие ошибки чаще всего допускают при статистической обработке данных? Как их избежать?
Практические задания (могут быть предложены на зачёте)
Постройте гистограмму и box-plot для заданного набора данных. Сделайте выводы о распределении признака.
Рассчитайте коэффициент корреляции Пирсона между двумя переменными. Интерпретируйте результат.
Выполните t-тест для сравнения средних двух независимых групп. Сформулируйте выводы.
Постройте линейную регрессионную модель по заданным данным. Оцените её качество (R2) и сделайте прогноз для нового значения X.

Критерии оценки на зачёте:
«Зачтено»: аспирант демонстрирует глубокое понимание теоретических основ, уверенно оперирует терминами, правильно решает практические задачи, аргументирует выводы.
«Не зачтено»: аспирант не раскрывает суть вопросов, допускает грубые ошибки в теории и практике, не может интерпретировать результаты анализа.

4. Учебно-методическое и информационное обеспечение дисциплины.
4.1. Основная литература
1. Афанасьев В.Н., Еремеева Н.С., Лебедева Т.В. Статистическая методология в научных исследованиях: учебное пособие для аспирантов. Оренбург: ОГУ, 2017. — 245 с.
2. Боровиков В.П. STATISTICA: Статистический анализ и обработка данных в среде Windows. М.: Филинъ, 1998.
3. Елисеева И.И., Юзбашев М.М. Общая теория статистики. М.: Финансы и статистика, 1999.
4. Юнкеров В.И., Григорьев С.Г. Математико-статистическая обработка данных медицинских исследований.
5. Шпаков П.С., Попов В.Н. Статистическая обработка экспериментальных данных: учебное пособие. М.: Недра, 2003.

Дополнительная литература
6. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. М.: Финансы и статистика, 1998.
7. Кендалл М., Стьюарт А. Статистические выводы и связи. М.: Наука, 1973.
8. Лакин Г.Ф. Биометрия: учебное пособие для студентов биологических специальностей. 4-е изд., перераб. и доп. М.: Высшая школа, 1990. — 351 с.
9. Медик В.А., Токмачев М.С., Фишман Б.Б. Статистика в медицине и биологии: в 2 томах.
10. Плохинский Н.А. Биометрия. М.: Издательство МГУ, 1970. — 367 с.

Сидоренко Е.В. Методы математической обработки в психологии.
Ермолаев О.Ю. Математическая статистика для психологов.
Кричевец А.Н., Шикин Е.В., Дьячков А.Г. Математика для психологов.
Малета Ю.С., Тарасов В.В. Непараметрические методы статистического анализа в биологии и медицине.
Доспехов Б.А. Методика полевого опыта (с основами статистической обработки результатов исследований).

Цифровые ресурсы и программное обеспечение
Официальные руководства к пакетам статистического ПО:
R: официальный сайт r-project.org — документация, учебники, пакеты для анализа.
SPSS: IBM SPSS Statistics Base User's Guide — руководство пользователя с примерами.
Python (библиотеки pandas, scipy, statsmodels, seaborn) — официальная документация и туториалы.
Онлайн-курсы и лекции:
Открытые курсы по статистике на Coursera, Stepik (темы: «Введение в статистику», «Анализ данных»).
Лекции и вебинары по работе с R и Python для статистического анализа.

Материально-техническое и программное обеспечение дисциплины (модуля)
5.1. Материально-техническое обеспечение

Вид аудитории	Технические средства и оборудование
Лекционная учебная аудитория	Набор демонстрационного мультимедийного оборудования для демонстрации учебных материалов. Персональный компьютер для преподавателя с доступом в Интернет.
Помещения для проведения занятий семинарского типа	Маркерная доска. Персональный компьютер для преподавателя с доступом в Интернет. Флипчарты для групповой работы обучающихся.
Компьютерный класс	Персональный компьютер для преподавателя с доступом в Интернет. Мультимедийное оборудование для демонстрации учебных материалов. Персональные компьютеры для обучающихся с доступом в Интернет. Программное обеспечение согласно нижеприведенному списку.

5.2. Перечень программного обеспечения
Программное обеспечение доступное для использования в АНОО ВО «Университет «Сириус».
Офисный пакет (LibreOffice) либо облачный офис (Яндекс. Документы).